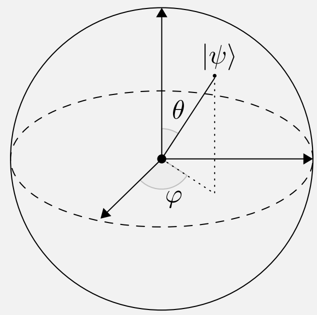
Solutions to Quantum Computation and Quantum Information



Arnab Adhikary, Rio Weil

This document was typeset on July 17, 2023

Introduction:

This document is a (work in progress) collection of comprehensive solutions to the exercises in Nielsen and Chuang's "Quantum Computation and Quantum Information". Each solution has the involved concepts (and hence rough pre-requisite knowledge) necessary for the problem in addition to the solution. Some problems may contain additional remarks about implications. Any problems denoted as (Research) are left as exercises to the reader. Starred exercises are considered to be more difficult (difficulty is assumed for the problems at the end of the chapter).

Contents

1 Introduction and overview

Exercise 1.1: Probabilistic Classical Algorithm

Suppose that the problem is not to distinguish between the constant and balanced functions with certainty, but rather, with some probability of error $\epsilon < 1/2$. What is the performance of the best classical algorithm for this problem?

Solution

Concepts Involved: Deutsch's Problem, Probability.
Recall that a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is said to be *balanced* if $f(x) = 1$ for exactly half of all possible 2^n values of x.

A single evaluation tells us no information about whether f is constant or balanced, so our success rate/error rate after a single evaluation is $\epsilon = \frac{1}{2}$ (random guessing!). Therefore, consider the case where we do two evaluations. If we obtain two different results, then we immediately conclude that f is balanced. Suppose instead that we obtain two results that are the same. If f is balanced, then the probably that the first evaluation returns the given result is $\frac{1}{2}$, and the probability that the second evaluation returns the same result is $\frac{2^n/2-1}{2^n-1}$ (as there are $2^n/2$ of each result of 0 and 1, 2^n total results, $2^n/2-1$ of the given result left after the first evaluation, and 2^n-1 total uninvestigated cases after the first evaluation). Therefore, if f is balanced, this occurs with probability $\frac{1}{2} \cdot \frac{2^n/2-1}{2^n-1}$, which we can see is less than $\frac{1}{2}$ as:

$$2^n < 2^{n+1} \implies 2^n - 2 < 2^{n+1} - 2 \implies \frac{2^n/2-1}{2^n-1} < 1 \implies \frac{1}{2}\frac{2^n/2-1}{2^n-1} < \frac{1}{2}$$

Hence, if we get the same result in two evaluations, we can conclude that f is constant with error $\epsilon < \frac{1}{2}$. We conclude that only 2 evaluations are required for this algorithm. $\square$

Exercise 1.2

Explain how a device which, upon input of one of two non-orthogonal quantum states $|\psi\rangle$ or $|\varphi\rangle$ correctly identified the state, could be used to build a device which cloned the states $|\psi\rangle$ and $|\varphi\rangle$, in violation of the no-cloning theorem. Conversely, explain how a device for cloning could be used to distinguish non-orthogonal quantum states.

Solution

Concepts Involved: Quantum Distinguishability, Quantum Measurement.

Given access to a device which can distinguish non-orthogonal quantum states $|\psi\rangle, |\varphi\rangle$ (without measurement), we can then design a quantum circuit that would map $|\psi\rangle \mapsto |\varphi\rangle$ (or vise versa), allowing us to clone the states as we like.
Conversely, given a cloning device, we could clone $|\psi\rangle$ and $|\varphi\rangle$ an arbitrary number of times. Then, performing repeated measurements of the two states in different measurement bases, we would (given enough measurements) be able to distinguish the two states based on the measurement statistics (there will of course be some error ϵ based on probabilistic considerations, but given that we have access to as many measurements of the states as we like, we are able to make this error arbitrarily low). $\square$

Problem 1.1: (Feynman-Gates conversation)

Construct a friendly imaginary discussion of about 2000 words between Bill Gates and Richard Feynman, set in the present, on the future of computation (Comment: You might like to try waiting until you've heard the rest of the book before attempting this question. See 'History and further reading' below for pointers to one possible answer for this question).

Problem 1.2

What is the most significant discovery yet made in quantum computation and quantum information? Write an essay of about 2000 words to an educated lay audience about the discovery (Comment: As for the previous problem, you might like to try waiting until you've read the rest of the book before attempting this question.)

2 Introduction to quantum mechanics

Exercise 2.1: Linear dependence: example

Show that $(1, -1)$, $(1, 2)$ and $(2, 1)$ are linearly dependent.

Solution

Concepts Involved: Linear Algebra, Linear Independence/Dependence.

We observe that:

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1+1-2 \\ -1+2-1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

showing that the three vectors are linearly dependent by definition. Alternatively, we can apply theorem that states that for any vector space V with $\dim V = n$, any list of $m > n$ vectors in V will be linearly dependent (here, $V = \mathbb{R}^2, n = 2, m = 3$). $\square$

Exercise 2.2: Matrix representations: example

Suppose V is a vector space with basis vectors $|0\rangle$ and $|1\rangle$, and A is a linear operator from V to V such that $A|0\rangle = |1\rangle$ and $A|1\rangle = |0\rangle$. Give a matrix representation for A, with respect to the input basis $|0\rangle$, $|1\rangle$, and the output basis $|0\rangle, |1\rangle$. Find input and output bases which give rise to a different matrix representation of A.

Solution

Concepts Involved: Linear Algebra, Matrix Representation of Operators.

Identifying $|0\rangle \cong \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $|1\rangle \cong \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, we have that:

$$A = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix}$$

Using the given relations, we have that:

$$A|0\rangle = 0|0\rangle + 1|1\rangle \implies \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \implies a_{00} = 0, a_{01} = 1$$
$$A|1\rangle = 1|0\rangle + 0|1\rangle \implies \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \implies a_{10} = 1, a_{11} = 0$$

Therefore with respect to the input basis $\{|0\rangle, |1\rangle\}$ and output basis $\{|0\rangle, |1\rangle\}$, A has matrix representation:

$$A \cong \begin{array}{c} \\ |0\rangle \\ |1\rangle \end{array} \begin{array}{c} \begin{array}{cc} \langle 0| & \langle 1| \end{array} \\ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \end{array}$$

Suppose we instead choose the input and output basis to be $\left\{ |+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}, |-\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}} \right\}$. Identifying $|+\rangle \cong \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $|-\rangle \cong \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, we have:

$$A = \begin{bmatrix} a_{++} & a_{+-} \\ a_{-+} & a_{--} \end{bmatrix}$$

Using the linearity of A, we have that:

$$A\,|+\rangle = \frac{1}{\sqrt{2}} A(|0\rangle + |1\rangle) = \frac{1}{\sqrt{2}}(A\,|0\rangle + A\,|1\rangle) = \frac{1}{\sqrt{2}}(|1\rangle + |0\rangle) = |+\rangle$$

and:

$$A\,|-\rangle = \frac{1}{\sqrt{2}} A(|0\rangle - |1\rangle) = \frac{1}{\sqrt{2}}(A\,|0\rangle - A\,|1\rangle) = \frac{1}{\sqrt{2}}(|1\rangle - |0\rangle) = -\,|-\rangle\,,$$

which can be used to determine the matrix elements:

$$A\,|+\rangle = 1\,|+\rangle + 0\,|0\rangle \implies \begin{bmatrix} a_{++} & a_{+-} \\ a_{-+} & a_{--} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \implies a_{++} = 1, a_{+-} = 0$$

$$A\,|-\rangle = 0\,|+\rangle - 1\,|-\rangle \implies \begin{bmatrix} a_{++} & a_{+-} \\ a_{-+} & a_{--} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \implies a_{-+} = 0, a_{--} = -1$$

Therefore with respect to the input basis $\{ |+\rangle, |-\rangle \}$ and output basis $\{ |+\rangle, |-\rangle \}$, A has matrix representation:

$$A \cong \begin{array}{c} \\ |+\rangle \\ |-\rangle \end{array} \begin{array}{c} \begin{array}{cc} \langle +| & \langle -| \end{array} \\ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \end{array}$$

□

Remark: If we choose the input and output bases to be different, we can even represent the A operator as an identity matrix. Specifically, if the input basis to be chosen to be $\{ |0\rangle, |1\rangle \}$ and output basis as $\{ |1\rangle, |0\rangle \}$, the matrix representation of A looks like:

$$A \cong \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Exercise 2.3: Matrix representation for operator products

Suppose A is a linear operator from vector space V to vector space W, and B is a linear operator from vector space W to vector space X. Let $|v_i\rangle, |w_j\rangle, |x_k\rangle$ be bases for the vector spaces V, W and X respectively. Show that the matrix representation for the linear transformation BA is the matrix product of the matrix representations for B and A, with respect to the appropriate bases.

Solution

Concepts Involved: Linear Algebra, Matrix Representation of Operators.

Taking the matrix of representations of A and B to the appropriate bases $|v_i\rangle, |w_j\rangle, |x_k\rangle$ of V, W and X, we have that:

$$A|v_j\rangle = \sum_i A_{ij}|w_i\rangle, \quad B|w_i\rangle = \sum_k B_{ki}|x_k\rangle$$

Hence, looking at $BA : V \mapsto X$, we have that:

$$\begin{aligned}
BA|v_j\rangle &= B(A|v_j\rangle) \\
&= B\left(\sum_i A_{ij}|w_i\rangle\right) \\
&= \sum_i A_{ij}B|w_i\rangle \\
&= \sum_i A_{ij}\left(\sum_k B_{ki}|x_k\rangle\right) \\
&= \sum_k \sum_i B_{ki}A_{ij}|x_k\rangle \\
&= \sum_k (BA)_{kj}|x_k\rangle
\end{aligned}$$

which shows that the matrix representation of BA is indeed the matrix product of the representations of B and A. $\square$

Exercise 2.4: Matrix representation for identity

Show that the identity operator on a vector space V has a matrix representation which is one along the diagonal and zero everywhere else, if the matrix is taken with respect to the same input and output bases. This matrix is known as the *identity matrix*.

Solution

Concepts Involved: Linear Algebra, Matrix Representation of Operators.

Let V be a vector space and $|v_i\rangle$ be a basis of V. Let $A : V \mapsto V$ be a linear operator, and let its matrix representation taken to be respect to $|v_i\rangle$ as the input and output basis. We then have that for each

$i \in \{1, \dots, n\}$:

$$A|v_i\rangle = 1|v_i\rangle + \sum_{j \neq i} 0|v_j\rangle = \sum_j \delta_{ij}|v_j\rangle$$

From which we obtain that A has the matrix representation:

$$A \cong \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

$\square$

Exercise 2.5

Verify that $(\cdot, \cdot)$ just defined is an inner product on $\mathbb{C}^n$.

Solution

Concepts Involved: Linear Algebra, Inner Products.
Recall that on $\mathbb{C}^n$, $(\cdot, \cdot)$ was defined as:

$$((y_1, \dots, y_n), (z_1, \dots, z_n)) \equiv \sum_i y_i^* z_i = [y_1^* \dots y_n^*] \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}.$$

Furthermore, recall the three conditions for the function $(\cdot, \cdot) : V \times V \mapsto \mathbb{C}$ to be considered an inner product:

(1) $(\cdot, \cdot)$ is linear in the second argument.

(2) $(|v\rangle, |w\rangle) = (|w\rangle, |v\rangle)^*$.

(3) $(|v\rangle, |v\rangle) \geq 0$ with equality if and only if $|v\rangle = \mathbf{0}$.

We check that $(\cdot, \cdot) : \mathbb{C}^n \times \mathbb{C}^n \mapsto \mathbb{C}$ satisfies the three conditions:

(1) We see that:

$$\begin{aligned} ((y_1, \dots, y_n), \sum_k \lambda_k (z_1, \dots, z_n)_k) &= \sum_i y_i^* \sum_k \lambda_k z_{i_k} \\ &= \sum_k \lambda_k \sum_i y_i^* z_{i_k} \\ &= \sum_k \lambda_k ((y_1, \dots, y_n), (z_1, \dots, z_n)_k) \end{aligned}$$

(2) We have:

$$\begin{aligned}((y_1, \dots, y_n), (z_1, \dots, z_n)) &= \sum_i y_i^* z_i \\ &= \sum_i (y_i z_i^*)^* \\ &= \left(\sum_i z_i^* y_i\right)^* \\ &= \left(((z_1, \dots, z_n), (y_1, \dots, y_n))\right)^*\end{aligned}$$

(3) We observe for $\mathbf{0} = (0, \dots 0)$:

$$(\mathbf{0}, \mathbf{0}) = \sum_i 0 \cdot 0 = 0$$

For $\mathbf{y} = (y_1, \dots, y_n) \neq 0$ we have that at least one y_i (say, y_j) is nonzero, and hence:

$$((y_1, \dots, y_n), (y_1, \dots, y_n)) = \sum_i y_i^2 \geq y_j^2 > 0$$

which proves the claim.

□

Exercise 2.6

Show that any inner product $(\cdot, \cdot)$ is conjugate-linear in the first argument,

$$\left(\sum_i \lambda_i \left|w_i\right\rangle, \left|v\right\rangle\right) = \sum_i \lambda_i^* (\left|w_i\right\rangle, \left|v\right\rangle).$$

Solution

Concepts Involved: Linear Algebra, Inner Products

Applying properties (2) (conjugate symmetry), (1) (linearity in second argument), and (2) (again) in succession, we have that:

$$\begin{aligned}
\left(\sum_i \lambda_i \left|w_i\right\rangle, \left|v\right\rangle\right) &= \left(|v\rangle, \sum_i \lambda_i \left|w_i\right\rangle\right)^* \\
&= \left(\sum_i \lambda_i (|v_i\rangle, |w_i\rangle)\right)^* \\
&= \sum_i \lambda_i^* (|v_i\rangle, |w_i\rangle)^* \\
&= \sum_i \lambda_i^* (|w_i\rangle, |v_i\rangle)
\end{aligned}$$

□

Exercise 2.7

Verify that $|w\rangle = (1, 1)$ and $|v\rangle = (1, -1)$ are orthogonal. What are the normalized forms of these vectors?

Solution

Concepts Involved: Linear Algebra, Inner Products, Orthogonality, Normalization
Recall that two vectors $|v\rangle, |w\rangle$ are orthogonal if $\langle v|w\rangle = 0$, and the norm of $|v\rangle$ is given by $\left\||v\rangle\right\| = \sqrt{\langle v|v\rangle}$.

First we show the two vectors are orthogonal:

$$\langle w|v\rangle = 1 \cdot 1 + 1 \cdot (-1) = 0$$

The norms of $\left|w\right\rangle, \left|v\right\rangle$ are given by:

$$\begin{aligned}
\left\||w\rangle\right\| &= \sqrt{\langle w|w\rangle} = \sqrt{1^2 + 1^2} = \sqrt{2}, \\
\left\||c\rangle\right\| &= \sqrt{\langle v|v\rangle} = \sqrt{1^2 + (-1)^2} = \sqrt{2}
\end{aligned}$$

So the normalized forms of the vectors are:

$$\begin{aligned}
\frac{|w\rangle}{\left\||w\rangle\right\|} &= \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \\
\frac{|v\rangle}{\left\||v\rangle\right\|} &= \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix}
\end{aligned}$$

□

Exercise 2.8

Verify that the Gram-Schmidt procedure produces and orthonormal basis for V.

Solution

Concepts Involved: Linear Algebra, Linear Independence, Bases, Inner Products, Orthogonality, Normalization, Gram-Schmidt Procedure, Induction.

Recall that given $|w_1\rangle, \dots, |w_d\rangle$ as a basis set for a vector space V, the Gram-Schmidt procedure constructs a basis set $|v_1\rangle, \dots, |v_d\rangle$ by defining $|v_1\rangle \equiv |w_1\rangle / \big\||w_1\rangle\big\|$ and then defining $|v_{k+1}\rangle$ inductively for $1 \leq k \leq d-1$ as:

$$|v_{k+1}\rangle \equiv \frac{|w_{k+1}\rangle - \sum_{i=1}^{k} \langle v_i|w_{k+1}\rangle |v_i\rangle}{\Big\||w_{k+1}\rangle - \sum_{i=1}^{k} \langle v_i|w_{k+1}\rangle |v_i\rangle\Big\|}$$

It is evident that each of the $|v_j\rangle$ have unit norm as they are defined in normalized form. It therefore suffices to show that each of the $|v_1\rangle, \dots, |v_d\rangle$ are orthogonal to each other, and that this set of vectors forms a basis of V. We proceed by induction. For $k = 1$, we have that:

$$|v_2\rangle = \frac{|w_2\rangle - \langle v_1|w_2\rangle |v_1\rangle}{\big\||w_2\rangle + \langle v_1|w_2\rangle |v_1\rangle\big\|}$$

Therefore:

$$\langle v_1|v_2\rangle = \frac{\langle v_1|w_2\rangle - \langle v_1|w_2\rangle \langle v_1|v_1\rangle}{\big\||w_2\rangle + \langle v_1|w_2\rangle |v_1\rangle\big\|} = \frac{\langle v_1|w_2\rangle - \langle v_1|w_2\rangle}{\big\||w_2\rangle + \langle v_1|w_2\rangle |v_1\rangle\big\|} = 0$$

so the two vectors are orthogonal. Furthermore, they are linearly independent; if they were linearly dependent, we could write $|v_1\rangle = \lambda |v_2\rangle$ for some $\lambda \in \mathbb{C}$, but then multiplying both sides by $\langle v_1|$ we get:

$$\langle v_1|v_1\rangle = \lambda \langle v_1|v_2\rangle \implies 1 = 0$$

which is a contradiction. This concludes the base case. For the inductive step, let $k \geq 1$ and suppose that $|v_1\rangle, \dots, |v_k\rangle$ are orthogonal and linearly independent. We then have that:

$$|v_{k+1}\rangle = \frac{|w_{k+1}\rangle - \sum_{i=1}^{k} \langle v_i|w_{k+1}\rangle |v_i\rangle}{\Big\||w_{k+1}\rangle - \sum_{i=1}^{k} \langle v_i|w_{k+1}\rangle |v_i\rangle\Big\|}$$

Then for any $j \in \{1, \dots k\}$, we have that:

$$\langle v_j|v_{k+1}\rangle = \frac{\langle v_j|w_{k+1}\rangle - \sum_{i=1}^{k} \langle v_i|w_{k+1}\rangle \langle v_j|v_j\rangle |v_i\rangle}{\Big\||w_{k+1}\rangle - \sum_{i=1}^{k} \langle v_i|w_{k+1}\rangle |v_i\rangle\Big\|} = \frac{\langle v_j|w_{k+1}\rangle - \langle v_j|w_{k+1}\rangle}{\Big\||w_{k+1}\rangle - \sum_{i=1}^{k} \langle v_i|w_{k+1}\rangle |v_i\rangle\Big\|} = 0$$

where in the second equality we use the fact that $\big\langle v_j\big|v_i\big\rangle = \delta_{ij}$ for $i, j \in \{1, \dots k\}$ by the inductive hypothesis. We therefore find that $|v_{k+1}\rangle$ is orthogonal to all of $|v_1\rangle, \dots, |v_k\rangle$. Furthermore, $|v_1\rangle, \dots, |v_k\rangle, |v_{k+1}\rangle$ is lienarly independent. Suppose for the sake of contradiction that this was false. Then, there would exist $\lambda_1, \dots \lambda_k$ not all nonzero such that:

$$\lambda_1|v_1\rangle + \dots + \lambda_k|v_k\rangle = |v_{k+1}\rangle$$

but then multiplying both sides by $\langle v_{k+1}|$ we have that:

$$\lambda_1 \langle v_{k+1}|v_1\rangle + \ldots + \lambda_k \langle v_{k+1}|v_k\rangle = \langle v_{k+1}|v_{k+1}\rangle \implies 0 = 1$$

by orthonormality. This gives a contradiction, and hence $|v_1\rangle, \ldots, |v_k\rangle, |v_{k+1}\rangle$ are linearly independent, finishing the inductive step. Therefore, $|v_1\rangle, \ldots, |v_d\rangle$ is an orthonormal list of vectors which is linearly independent. Since $|w_1\rangle, \ldots, |w_d\rangle$ is a basis for V, then V has dimension d. Hence, $|v_1\rangle, \ldots, |v_d\rangle$ being a linearly independent list of d vectors in V is a basis of V. We conclude that it is an orthonormal basis of V, as claimed. □

Exercise 2.9: Pauli operators and the outer product

The Pauli matrices (Figure 2.2 on page 65) can be considered as operators with respect to an orthonormal basis $|0\rangle, |1\rangle$ for a two-dimensional Hilbert space. Express each of the Pauli operators in the outer product notation.

Solution

Concepts Involved: Linear Algebra, Matrix Representation of Operators, Outer Products.
Recall that if A has matrix representation:

$$A \cong \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix}$$

with respect to $|0\rangle, |1\rangle$ as the input/output bases, then we can express A in outer product notation as:

$$A = a_{00}\,|0\rangle\langle 0| + a_{01}\,|0\rangle\langle 1| + a_{10}\,|1\rangle\langle 0| + a_{11}\,|1\rangle\langle 1|$$

Furthermore, recall the representation of the Pauli matrices with respect to the orthonormal basis $|0\rangle, |1\rangle$:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

We immediately see that:

$$\begin{aligned} I &= |0\rangle\langle 0| + |1\rangle\langle 1| \\ X &= |0\rangle\langle 1| + |1\rangle\langle 0| \\ Y &= -i\,|0\rangle\langle 1| + i\,|1\rangle\langle 0| \\ Z &= |0\rangle\langle 0| - |1\rangle\langle 1| \end{aligned}$$

□

Exercise 2.10

Suppose $|v_i\rangle$ is an orthonormal basis for an inner product space V. What is the matrix representation for the operator $|v_j\rangle\langle v_j|$, with respect to the $|v_i\rangle$ basis?

Solution

Concepts Involved: Linear Algebra, Matrix Representation of Operators, Outer Products.

The matrix representation of $|v_j\rangle\langle v_j|$ with respect to the $|v_i\rangle$ basis is a matrix with 1 in the jth column and row (i.e. the (j,j)th entry in the matrix) and 0 everywhere else. $\square$

Exercise 2.11

Find the eigenvectors, eigenvalues, and diagonal representations of the Pauli matrices X, Y and Z.

Solution

Concepts Involved: Linear Algebra, Eigenvalues, Eigenvectors, Diagonalization.

Given an operator A on a vector space V, recall that an eigenvector $|v\rangle$ of A and its corresponding eigenvalue λ are defined by:

$$A|v\rangle = \lambda|v\rangle$$

Furthermore, recall the diagonal representation of A is given by

$$A = \sum_i \lambda_i \, |i\rangle\langle i|$$

Where $|i\rangle$ form an orthonormal set of eigenvectors for A, and λ_i are the corresponding eigenvalues.

We start with X. Solving for the eigenvalues, we have:

$$\det(X - I\lambda) = 0 \implies \det\begin{bmatrix} -\lambda & 1 \\ 1 & -\lambda \end{bmatrix} = 0 \implies \lambda^2 - 1 = 0$$

From which we obtain $\lambda_1 = 1, \lambda_2 = -1$. Solving for the eigenvectors, we then have that:

$$(X - I\lambda_1)|v_1\rangle = \mathbf{0} \implies \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies v_{11} = 1, v_{12} = 1$$

$$(X - I\lambda_2)|v_2\rangle = \mathbf{0} \implies \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies v_{21} = 1, v_{22} = -1$$

Hence we find that $|v_1\rangle = |0\rangle + |1\rangle$, $|v_2\rangle = |0\rangle - |1\rangle$. Normalizing these eigenvectors (Also see Exercise 2.7), we divide by $\||v_1\rangle\| = \||v_2\rangle\| = \sqrt{2}$, giving us:

$$|v_1\rangle = |+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad |v_2\rangle = |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

The diagonal representation of X is then given by:

$$X = \lambda_1 \, |v_1\rangle\langle v_1| + \lambda_2 \, |v_2\rangle\langle v_2| = \, |+\rangle\langle +| - \, |-\rangle\langle -|$$

We do the same for Y. Solving for the eigenvalues:

$$\det(A - I\lambda) = 0 \implies \det\begin{bmatrix} -\lambda & -i \\ i & -\lambda \end{bmatrix} = 0 \implies \lambda^2 - 1 = 0$$

From which we obtain $\lambda_1 = 1, \lambda_2 = -1$. Solving for the eigenvectors, we then have that:

$$(Y - I\lambda_1)|v_1\rangle = \mathbf{0} \implies \begin{bmatrix} -1 & -i \\ i & -1 \end{bmatrix}\begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies v_{11} = 1, v_{12} = i$$
$$(Y - I\lambda_2)|v_2\rangle = \mathbf{0} \implies \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}\begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies v_{21} = 1, v_{22} = -i$$

We therefore have that $|v_1\rangle = |0\rangle + i|1\rangle, |v_2\rangle = |0\rangle - i|1\rangle$. Normalizing by dividing by $\||v_1\rangle\| = \||v_2\rangle\|$, we obtain that:

$$|v_1\rangle = |y_+\rangle = \frac{|0\rangle + i|1\rangle}{\sqrt{2}}, \quad |v_2\rangle = |y_-\rangle = \frac{|0\rangle - i|1\rangle}{\sqrt{2}}.$$

The diagonal representation of Y is then given by:

$$Y = |y_+\rangle\langle y_+| - |y_-\rangle\langle y_-|$$

For Z, the process is again the same. We give the results and omit the details:

$$\lambda_1 = 1, |v_1\rangle = |0\rangle \quad \lambda_2 = -1, |v_2\rangle = |1\rangle$$

$$Z = |0\rangle\langle 0| - |1\rangle\langle 1|$$

□

Exercise 2.12

Prove that the matrix

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

is not diagonalizable.

Solution

Concepts Involved: Linear Algebra, Eigenvalues, Eigenvectors, Diagonalization.

Solving for the eigenvalues of the matrix, we have:

$$\det\begin{bmatrix} 1-\lambda & 0 \\ 1 & 1-\lambda \end{bmatrix} = 0 \implies (1-\lambda)^2 = 0 \implies \lambda_1, \lambda_2 = 1$$

But since the eigenvalue 1 is degenerate, the matrix only has one eigenvector; it therefore cannot be diagonalized. $\square$

Exercise 2.13

If $|w\rangle$ and $|v\rangle$ are any two vectors, show that $(|w\rangle\langle v|)^\dagger = |v\rangle\langle w|$.

Solution

Concepts Involved: Linear Algebra, Adjoints.

We observe that:

$$\begin{aligned}((|w\rangle\langle v|)^\dagger|x\rangle, |y\rangle) = (|x\rangle, (|w\rangle\langle v|)|y\rangle) = (|x\rangle, \langle v|y\rangle\,|w\rangle) &= \langle x|\,\langle v|y\rangle\,|w\rangle \\ &= \langle x|w\rangle\,\langle v|y\rangle \\ &= \langle x|w\rangle\,(|v\rangle, |y\rangle) \\ &= (\langle x|w\rangle^*\,|v\rangle, |y\rangle) \\ &= (\langle w|x\rangle\,|v\rangle, |y\rangle) \\ &= ((|v\rangle\langle w|)|x\rangle, |y\rangle)\end{aligned}$$

Where in the third-to last equality we use the conjugate linearity in the first argument (see Exercise 2.6) and in the second-to last equality we use that $\langle a|b\rangle^* = \langle b|a\rangle$. Comparing the first and last expressions, we conclude that $(|w\rangle\langle v|)^\dagger = |v\rangle\langle w|$. $\square$

Exercise 2.14: Anti-linearity of the adjoint

Show that the adjoint operator is anti-linear,

$$\left(\sum_i a_i A_i\right)^\dagger = \sum_i a_i^* A_i^\dagger.$$

Solution

Concepts Involved: Linear Algebra, Adjoints.

We observe that:

$$\begin{aligned}\left(\left(\sum_i a_i A_i\right)^\dagger |a\rangle, |b\rangle\right) = \left(|a\rangle, \sum_i a_i A_i|b\rangle\right) &= \sum_i a_i\,(|a\rangle, A_i|b\rangle) \\ &= \sum_i a_i\left(A_i^\dagger|a\rangle, |b\rangle\right) \\ &= \left(\sum_i a_i^* A_i^\dagger|a\rangle, |b\rangle\right)\end{aligned}$$

where we invoke the definition of the adjoint in the first and third equalities, the linearity in the second argument in the second equality, and the conjugate linearity in the first argument in the last equality. The claim is proven by comparing the first and last expressions. □

Exercise 2.15

Show that $(A^\dagger)^\dagger = A$.

Solution

Concepts Involved: Linear Algebra, Adjoints

Applying the definition of the Adjoint twice (and using the conjugate symmetry of the inner product) we have that:

$$((A^\dagger)^\dagger|a\rangle, |b\rangle) = (|a\rangle, A^\dagger|b\rangle) = (A^\dagger|b\rangle, |a\rangle)^* = (|b\rangle, A|a\rangle)^* = \left((A|a\rangle, |b\rangle)^*\right)^* = (A|a\rangle, |b\rangle).$$

The claim follows by comparison of the first and last expressions. □

Exercise 2.16

Show that any projector P satisfies the equation $P^2 = P$.

Solution

Concepts Involved: Linear Algebra, Projectors.

Let $|1\rangle, \ldots, |k\rangle$ be an orthonormal basis for the subspace W of V. Then, using the definition of the projector onto W, we have that:

$$P^2 = P \cdot P = \left(\sum_{i=1}^k |i\rangle\langle i|\right)\left(\sum_{i'=1}^k |i'\rangle\langle i'|\right) = \sum_{i=1}^k \sum_{i'=1}^k |i\rangle\langle i|i'\rangle\langle i'| = \sum_{i=1}^k \sum_{i'=1}^k |i\rangle\delta_{ii'}\langle i'| = \sum_{i=1}^k |i\rangle\langle i| = P$$

where in the fourth/fifth equality we use the orthonormality of the basis to collapse the double sum. □

Exercise 2.17

Show that a normal matrix is Hermitian if and only if it has real eigenvalues.

Solution

Concepts Involved: Linear Algebra, Hermitian Operators, Normal Operators, Spectral Decomposition.

$\boxed{\Longrightarrow}$ Let A be a Normal and Hermitian matrix. Then, it has a diagonal representation $A = \sum_i \lambda_i |i\rangle\langle i|$ where $|i\rangle$ is an orthonormal basis for V and each $|i\rangle$ is an eigenvector of A with eigenvalue λ_i. By the

Hermicity of A, we have that $A = A^\dagger$. Therefore, we have that:

$$A^\dagger = \left(\sum_i \lambda_i \left|i\middle\rangle\middle\langle i\right|\right)^\dagger = \sum_i \lambda_i^* \left|i\middle\rangle\middle\langle i\right| = A = \sum_i \lambda_i \left|i\middle\rangle\middle\langle i\right|$$

where we use the results of Exercises 2.13 and 2.14 in the second equality. Comparing the third and last expressions, we have that $\lambda_i = \lambda_i*$ and hence the eigenvalues are real.

$\boxed{\Longleftarrow}$ Let A be a Normal matrix with real eigenvalues. Then, A has diagonal representation $A = \sum_i \lambda_i \left|i\middle\rangle\middle\langle i\right|$ where λ_i are all real. We therefore have that:

$$A^\dagger = \left(\sum_i \lambda_i \left|i\middle\rangle\middle\langle i\right|\right)^\dagger = \lambda_i^* \left|i\middle\rangle\middle\langle i\right| = \sum_i \lambda_i \left|i\middle\rangle\middle\langle i\right| = A$$

where in the third equality we use that $\lambda_i^* = \lambda_i$. We conclude that A is Hermitian. $\square$

Exercise 2.18

Show that all eigenvalues of a unitary matrix have modulus 1, that is, can be written in the form $e^{i\theta}$ for some real θ.

Solution

Concepts Involved: Linear Algebra, Unitary Operators, Spectral Decomposition

---

Let U be a unitary matrix. It is then normal as $U^\dagger U = U^\dagger U = I$. It therefore has spectral decomposition $U = \sum_i \lambda_i \left|i\middle\rangle\middle\langle i\right|$ where $|i\rangle$ is an orthonormal basis of V, and $|i\rangle$ are the eigenvectors of U with eigenvalues λ_i. We then have that:

$$\begin{aligned}
UU^\dagger = I &\implies \left(\sum_i \lambda_i \left|i\middle\rangle\middle\langle i\right|\right)\left(\sum_{i'} \lambda_{i'} \left|i'\middle\rangle\middle\langle i'\right|\right)^\dagger = I \\
&\implies \left(\sum_i \lambda_i \left|i\middle\rangle\middle\langle i\right|\right)\left(\sum_{i'} \lambda_{i'}^* \left|i'\middle\rangle\middle\langle i'\right|\right) = I \\
&\implies \sum_i \sum_{i'} \lambda_i \lambda_{i'} |i\rangle \langle i|i'\rangle \langle i'| = I \\
&\implies \sum_i \sum_{i'} \lambda_i \lambda_{i'}^* |i\rangle \delta_{ii'} \langle i'| = I \\
&\implies \sum_i \lambda_i \lambda_{i'} \left|i\middle\rangle\middle\langle i\right| = I \\
&\implies \sum_i |\lambda_i|^2 \left|i\middle\rangle\middle\langle i\right| = \sum_i 1 \left|i\middle\rangle\middle\langle i\right|
\end{aligned}$$

From which we obtain that $|\lambda_i|^2 = 1$, and hence $|\lambda_i| = 1$, proving the claim. $\square$

Exercise 2.19: Pauli matrices: Hermitian and unitary

Show that the Pauli matrices are Hermitian and unitary.

Solution

Concepts Involved: Linear Algebra, Hermitian Matrices, Unitary Matrices

We check I, X, Y, Z in turn.

$$I^\dagger = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^T\right)^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

$$I^\dagger I = II = I$$

$$X^\dagger = \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^T\right)^* = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^* = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = X$$

$$X^\dagger X = XX = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

$$Y^\dagger = \left(\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}^T\right)^* = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}^* = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = Y$$

$$Y^\dagger Y = YY = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$Z^\dagger = \left(\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^T\right)^* = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^* = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = Z$$

$$Z^\dagger Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

$\square$

Exercise 2.20: Basis changes

Suppose A' and A'' are matrix representations of an operator A on a vector space A on a vector space V with respect to two different orthonormal bases, $|v_i\rangle$ and $|w_i\rangle$. Then the elements of A' and A'' are $A'_{ij} = \langle v_i|A|v_j\rangle$ and $A''_{ij} = \langle w_i|A|w_j\rangle$. Characterize the relationship between A' and A''.

Solution

Concepts Involved: Linear Algebra, Matrix Representations of Operators, Completeness Relation

---

Using the completeness relation twice, we get:

$$\begin{aligned} A'_{ij} = \langle v_i|A|v_j\rangle = \langle v_i|IAI|v_j\rangle &= \langle v_i| \left(\sum_{i'} |w_{i'}\rangle\langle w_{i'}|\right) A \left(\sum_{j'} |w_{j'}\rangle\langle w_{j'}|\right) |v_j\rangle \\ &= \sum_{i'}\sum_{j'} \langle v_i|w_{i'}\rangle\langle w_{i'}|A|w_{j'}\rangle\langle w_{j'}|v_j\rangle \\ &= \sum_{i'}\sum_{j'} \langle v_i|w_{i'}\rangle A''_{ij}\langle w_{j'}|v_j\rangle \end{aligned}$$

□

Exercise 2.21

Repeat the proof of the spectral decomposition in Box 2.2 for the case when M is Hermitian, simplifying the proof wherever possible.

Solution

Concepts Involved: Linear Algebra, Hermitian Operators, Spectral Decomposition.

For the converse, we have that if M is diagonalizable, then it has a representation $M = \sum_i \lambda_i |i\rangle\langle i|$ where $|i\rangle$ is an orthonormal basis of V, and $|i\rangle$ are eigenvectors of M with associated eigenvalues of λ_i. We then have that:

$$M^\dagger = \left(\sum_i \lambda_i |i\rangle\langle i|\right)^\dagger = \sum_i \lambda_i^* |i\rangle\langle i| = \sum_i \lambda_i |i\rangle\langle i| = M$$

where in the second equality we apply the result of Exercise 2.13 and in the third equality we use that Hermitian matrices have real eigenvalues. For the forwards implication, we proceed by induction on the dimension d of V. The $d = 1$ case is trivial as M is already diagonal in any representation in this case. Let λ be an eigenvalue of M, P the projector onto the λ subspace, and Q the projector onto the orthogonal complement. Then $M = (P+Q)M(P+Q) = PMP + QMP + PMQ + QMQ$. Obviously $PMP = \lambda P$. Furthermore, $QMP = 0$, as M takes the subspace P into itself. We claim that $PMQ = 0$ also. To see this, we recognize that $(PMQ)^\dagger = Q^\dagger M^\dagger P^\dagger = QMP = 0$. and hence $PMQ = 0$. Thus $M = PMP + QMQ$. QMQ is normal, as $(QMQ)^\dagger = Q^\dagger M^\dagger Q^\dagger = QMQ$ (and Hermiticity implies that the operator is normal). By induction, QMQ is diagonal with respect to some orthonormal basis for the subspace Q, and PMP is already diagonal with respect to some orthonormal basis for P. It follows that $M = PMP + QMQ$ is diagonal with respect to some orthonormal basis for the total vector space. $\square$

Exercise 2.22

Prove that two eigenvectors of a Hermitian operator with different eigenvalues are necessarily orthogonal.

Solution

Concepts Involved: Linear Algebra, Eigenvalues, Eigenvectors, Hermitian Operators.

Let A be a Hermitian operator, and let $|v_1\rangle, |v_2\rangle$ be two eigenvectors of A with corresponding eigenvalues λ_1, λ_2 such that $\lambda_1 \neq \lambda_2$. We then have that:

$$\begin{aligned}
\langle v_1|A|v_2\rangle &= \langle v_1|\lambda_2|v_2\rangle = \lambda_2\langle v_1|v_2\rangle \\
\langle v_1|A|v_2\rangle &= \langle v_1|A^\dagger|v_2\rangle = \langle v_1|\lambda_1|v_2\rangle = \lambda_1\langle v_1|v_2\rangle
\end{aligned}$$

where we use the Hermiticity of A in the second line. Substracting the first line from the second, we have that:

$$0 = (\lambda_2 - \lambda_1)\langle v_1|v_2\rangle.$$

Since $\lambda_1 \neq \lambda_2$ by assumption, the only way this equality is satisfied is if $\langle v_1|v_2\rangle = 0$. Hence, $|v_1\rangle, |v_2\rangle$ are orthogonal. $\square$

Exercise 2.23

Show that the eigenvalues of a projector P are all either 0 or 1.

Solution

Concepts Involved: Linear Algebra, Eigenvalues, Eigenvectors, Projectors.

Let P be a projector, and $|v\rangle$ be an eigenvector of P with corresponding eigenvalue λ. From Exercise 2.16, we have that $P^2 = P$, and using this fact, we observe:

$$\begin{aligned}
&P|v\rangle = \lambda|v\rangle \\
&P|v\rangle = P^2|v\rangle = PP|v\rangle = P\lambda|v\rangle = \lambda P|v\rangle = \lambda^2|v\rangle.
\end{aligned}$$

Subtracting the first line from the second, we get:

$$\mathbf{0} = (\lambda^2 - \lambda)|v\rangle = \lambda(\lambda - 1)|v\rangle.$$

Since $|v\rangle$ is not the zero vector, we therefore obtain that either $\lambda = 0$ or $\lambda = 1$. $\square$

Exercise 2.24: Hermiticity of positive operator

Show that a positive operator is necessarily Hermitian. (Hint: Show that an arbitrary operator A can be written $A = B + iC$ where B and C are Hermitian.)

Solution

Concepts Involved: Linear Algebra, Hermitian Operators, Positive Operators

Let A be an operator. We first make the observation that we can write A as:

$$A = \frac{A}{2} + \frac{A}{2} + \frac{A^\dagger}{2} - \frac{A^\dagger}{2} = \frac{A + A^\dagger}{2} + i\frac{A - A^\dagger}{2i}.$$

So let $B = \frac{A+A^\dagger}{2}$ and $C = \frac{A-A^\dagger}{2i}$. B and C are Hermitian, as:

$$B^\dagger = \left(\frac{A + A^\dagger}{2}\right)^\dagger = \frac{A^\dagger + (A^\dagger)^\dagger}{2} = \frac{A^\dagger + A}{2} = B$$
$$C^\dagger = \left(\frac{A - A^\dagger}{2i}\right)^\dagger = \frac{A^\dagger - (A^\dagger)^\dagger}{-2i} = \frac{A - A^\dagger}{2i} = C$$

so we have hence proven that we can write $A = B + iC$ for hermitian B, C for any operator A. Now, assume that A is positive. We then have that for any vector $|v\rangle$:

$$\langle v|A|v\rangle \geq 0.$$

Using the identity derived above, we have that:

$$\langle v|B|v\rangle + i\langle v|C|v\rangle \geq 0.$$

The positivity forces $C = 0$. Therefore, $A = B$ and hence A is Hermitian. □

Exercise 2.25

Show that for any operator A, $A^\dagger A$ is positive.

Solution

Concepts Involved: Linear Algebra, Adjoints, Positive Operators

Let A be an operator. Let $|v\rangle$ be an arbitrary vector, and then we then have that:

$$\Big(|v\rangle, A^\dagger A|v\rangle\Big) = \Big((A^\dagger)^\dagger|v\rangle, A|v\rangle\Big) = \big(A|v\rangle, A|v\rangle\big).$$

By the property of inner products, the expression must be greater than zero. □

Exercise 2.26

Let $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. Write out $|\psi\rangle^{\otimes 2}$ and $|\psi\rangle^{\otimes 3}$ explicitly, both in terms of tensor products like $|0\rangle|1\rangle$ and using the Kronecker product.

Solution

Concepts Involved: Linear Algebra, Tensor Products, Kronecker Products.

Using the definition of the tensor product, we have:

$$|\psi\rangle^{\otimes 2} = \frac{|0\rangle|0\rangle + |0\rangle|1\rangle + |1\rangle|0\rangle + |1\rangle|1\rangle}{2} \cong \begin{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \\ \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

$$\begin{aligned} |\psi\rangle^{\otimes 3} &= \frac{|0\rangle|0\rangle|0\rangle + |0\rangle|0\rangle|1\rangle + |0\rangle|1\rangle|0\rangle + |0\rangle|1\rangle|1\rangle + |1\rangle|0\rangle|0\rangle + |1\rangle|0\rangle|1\rangle + |1\rangle|1\rangle|0\rangle + |1\rangle|1\rangle|1\rangle}{2\sqrt{2}} \\ &= |\psi\rangle \otimes |\psi\rangle^{\otimes 2} \cong \begin{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \\ \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} \end{bmatrix} \end{aligned}$$

$\square$

Exercise 2.27

Calculate the matrix representation of the tensor products of the Pauli operators (a) X and Z; (b) I and X; (c)X and I. Is the tensor product commutative?

Solution

Concepts Involved: Linear Algebra, Tensor Products, Kronecker Products.

Using the Kronecker product, we have:

(a)

$$X \otimes Z = \begin{bmatrix} 0Z & 1Z \\ 1Z & 0Z \end{bmatrix} = \begin{bmatrix} 0\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} & 1\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\ 1\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} & 0\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

(b)

$$I \otimes X = \begin{bmatrix} 1X & 0X \\ 0X & 1X \end{bmatrix} = \begin{bmatrix} 1\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} & 0\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ 0\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} & 1\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

(c)

$$X \otimes I = \begin{bmatrix} 0I & 1I \\ 1I & 0I \end{bmatrix} = \begin{bmatrix} 0\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & 1\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ 1\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & 0\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Comparing (b) and (c), we conclude that the tensor product is not commutative. □

Exercise 2.28

Show that the transpose, complex conjugation and adjoint operations distribute over the tensor product,

$$(A \otimes B)^* = A^* \otimes B^*; (A \otimes B)^T = A^T \otimes B^T; (A \otimes B)^\dagger = A^\dagger \otimes B^\dagger.$$

Solution

Concepts Involved: Linear Algebra, Adjoints, Tensor Products, Kronecker Products.

Using the Kronecker product representaton of $a \otimes B$, we have:

$$(A \otimes B)^* = \begin{bmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & A_{22}B & \dots & A_{2n}B \\ \vdots & \vdots & \vdots & \vdots \\ A_{m1}B & A_{m2}B & \dots & A_{mn}B \end{bmatrix}^* = \begin{bmatrix} A_{11}^*B^* & A_{12}^*B^* & \dots & A_{1n}^*B^* \\ A_{21}^*B^* & A_{22}^*B^* & \dots & A_{2n}^*B^* \\ \vdots & \vdots & \vdots & \vdots \\ A_{m1}^*B^* & A_{m2}^*B^* & \dots & A_{mn}^*B^* \end{bmatrix} = A^* \otimes B^*$$

$$(A \otimes B)^T = \begin{bmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & A_{22}B & \dots & A_{2n}B \\ \vdots & \vdots & \vdots & \vdots \\ A_{m1}B & A_{m2}B & \dots & A_{mn}B \end{bmatrix}^T = \begin{bmatrix} A_{11}B^T & A_{21}B^T & \dots & A_{n1}B^T \\ A_{12}B^T & A_{22}B & \dots & A_{n2}B^T \\ \vdots & \vdots & \vdots & \vdots \\ A_{1m}B^T & A_{2m}B^T & \dots & A_{nm}B^T \end{bmatrix} = A^T \otimes B^T.$$

The relation for the distributivity of the hermitian conjugate over the tensor product then follows from the former two relations:

$$(A \otimes B)^\dagger = ((A \otimes B)^T)^* = (A^T \otimes B^T)^* = (A^T)^* \otimes (B^T)^* = A^\dagger \otimes B^\dagger$$

□

Exercise 2.29

Show that the tensor product of two unitary operators is unitary.

Solution

Concepts Involved: Linear Algebra, Unitary Operators, Tensor Products

Suppose A, B are unitary. Then, $A^\dagger A = I$ and $B^\dagger B = I$. Using the result of the Exercise 2.28, we then have that:

$$(A \otimes B)^\dagger (A \otimes B) = (A^\dagger \otimes B^\dagger)(A \otimes B) = (A^\dagger A \otimes B^\dagger B) = I \otimes I$$

□

Exercise 2.30

Show that the tensor product of two Hermitian operators is Hermitian.

Solution

Concepts Involved: Linear Algebra, Hermitian Operators, Tensor Products

Suppose A, B are Hermitian. Then, $A^\dagger = A$ and $B^\dagger = B$. Then, using the result of Exercise 2.28, we have:

$$(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger = A \otimes B$$

□

Exercise 2.31

Show that the tensor product of two positive operators is positive.

Solution

Concepts Involved: Linear Algebra, Positive Operators

Suppose A, B are positive operators. We then have that $\langle v|A|v\rangle \geq 0$ and $\langle w|B|w\rangle \geq 0$. Therefore, for any $|v\rangle \otimes |w\rangle$:

$$\big(|v\rangle \otimes |w\rangle, A \otimes B(|v\rangle \otimes |w\rangle)\big) = \langle v|A|v\rangle\langle w|B|w\rangle \geq 0$$

$\square$

Exercise 2.32

Show that the tensor product of two projectors is a projector.

Solution

Concepts Involved: Linear Algebra, Projectors

Let P_1, P_2 be projectors. We then have that $P_1^2 = P_1$ and $P_2^2 = P_2$ by Exercise 2.16. Therefore:

$$(P_1 \otimes P_2)^2 = (P_1 \otimes P_2)(P_1 \otimes P_2) = P_1^2 \otimes P_2 = P_1 \otimes P_2$$

so $P_1 \otimes P_2$ is a projector. $\square$

Exercise 2.33

The Hadamard operator on one qubit may be written as

$$H = \frac{1}{\sqrt{2}}[(|0\rangle + |1\rangle)\langle 0| + (|0\rangle - |1\rangle)\langle 1|]$$

Show explicitly that the Hadamard transform on n qubits, $H^{\otimes n}$, may be written as

$$H^{\otimes n} = \frac{1}{\sqrt{2^n}}\sum_{x,y}(-1)^{x\cdot y}|x\rangle\langle y|$$

Write out an explicit matrix representation for $H^{\otimes 2}$

Solution

Concepts Involved: Linear algebra, Matrix Representation of Operators, Outer Products.

Looking at the form of the Hadamard operator on one qubit, we observe that:

$$H = \frac{1}{\sqrt{2}}\left[\,|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| - |1\rangle\langle 1|\right]$$

Hence:

$$H = \frac{1}{\sqrt{2}} \sum_{x,y} (-1)^{x \cdot y} |x\rangle\langle y|$$

Where x, y run over 0 and 1. Taking the n-fold tensor product of this expression, we get:

$$\begin{aligned} H^{\otimes} &= \frac{1}{\sqrt{2}} \sum_{x,y} (-1)^{x \cdot y} |x\rangle\langle y| \otimes \frac{1}{\sqrt{2}} \sum_{x,y} (-1)^{x \cdot y} |x\rangle\langle y| \otimes \ldots \otimes \frac{1}{\sqrt{2}} \sum_{x,y} (-1)^{x \cdot y} |x\rangle\langle y| \\ &= \frac{1}{\sqrt{2^n}} \sum_{\mathbf{x},\mathbf{y}} (-1)^{\mathbf{x} \cdot \mathbf{y}} |\mathbf{x}\rangle\langle \mathbf{y}| \end{aligned}$$

Where $\mathbf{x}, \mathbf{y}$ are length n-binary strings. This proves the claim.
Now explicitly writing $H^{\otimes 2}$, we have:

$$\begin{aligned} H^{\otimes 2} &= \frac{1}{\sqrt{2^2}} \sum_{\mathbf{x},\mathbf{y}} (-1)^{(\mathbf{x} \cdot \mathbf{y})} |\mathbf{x}\rangle\langle \mathbf{y}| \\ &\cong \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \end{aligned}$$

Note that here, $\mathbf{x}, \mathbf{y}$ are binary length 2 strings. The sum goes through all pairwise combinations of $\mathbf{x}, \mathbf{y} \in \{00, 01, 10, 11\}$. ☐

Remark: Sylvester's Construction gives an interesting recursive construction of Hadamard matrices. See `https://en.wikipedia.org/wiki/Hadamard_matrix`. Discussion on interesting (related) open problem concerning the maximal determinant of matrices consisting of entries of 1 and -1 can be found here `https://en.wikipedia.org/wiki/Hadamard%27s_maximal_determinant_problem`.

Exercise 2.34

Find the square root and logarithm of the matrix

$$\begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}$$

Solution

Concepts Involved: Linear Algebra, Spectral Decomposition, Operator Functions

We begin by diagonalizing the matrix (which we call A) as to be able to apply the definition of operator functions. By inspection, A is Hermitian as it is equal to its conjugate transpose, so the spectral decomposition exists. Solving for the eigenvalues, we consider the characterstic equation:

$$\det(A-\lambda I)=0 \implies \det\begin{bmatrix}4-\lambda & 3\\ 3 & 4-\lambda\end{bmatrix}=0 \implies (4-\lambda)^2-9=0 \implies \lambda^2-8\lambda+7=0$$

Using the quadratic equation, we get $\lambda_1=1, \lambda_2=7$. Using this to find the eigenvectors of the matrix, we have:

$$\begin{bmatrix}4-1 & 3\\ 3 & 4-1\end{bmatrix}\begin{bmatrix}v_{11}\\ v_{12}\end{bmatrix}=\mathbf{0} \implies v_{11}=1, v_{12}=-1$$

$$\begin{bmatrix}4-7 & 3\\ 3 & 4-7\end{bmatrix}\begin{bmatrix}v_{21}\\ v_{22}\end{bmatrix}=\mathbf{0} \implies v_{21}=1, v_{22}=1$$

Hence our normalized eigenvectors are:

$$|v_1\rangle=\begin{bmatrix}\frac{1}{\sqrt{2}}\\ \frac{-1}{\sqrt{2}}\end{bmatrix}, \quad |v_2\rangle=\begin{bmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{bmatrix}$$

Therefore the spectral composition of the matrix is given by:

$$A=1\,|v_1\rangle\langle v_1|+7\,|v_2\rangle\langle v_2|$$

Calculating the square root of A, we then have:

$$\sqrt{A}=\sqrt{1}\,|v_1\rangle\langle v_1|+\sqrt{7}\,|v_2\rangle\langle v_2|=\frac{1}{2}\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}+\frac{\sqrt{7}}{2}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}=\frac{1}{2}\begin{bmatrix}1+\sqrt{7} & -1+\sqrt{7}\\ -1+\sqrt{7} & 1+\sqrt{7}\end{bmatrix}.$$

Calculating the logarithm of A, we have:

$$\log(A)=\log(1)\,|v_1\rangle\langle v_1|+\log(7)\,|v_2\rangle\langle v_2|=\frac{\log(7)}{2}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$$

$\square$

Exercise 2.35: Exponential of Pauli matrices

Let $\mathbf{v}$ be any real, three-dimensional unit vector and θ a real number. Prove that

$$\exp(i\theta\mathbf{v}\cdot\boldsymbol{\sigma})=\cos(\theta)I+i\sin(\theta)\mathbf{v}\cdot\boldsymbol{\sigma}$$

Where $\mathbf{v}\cdot\boldsymbol{\sigma}\equiv\sum_{i=1}^{3}v_i\sigma_i$. This exercise is generalized in Problem 2.1 on page 117.

Solution

Concepts Involved: Linear Algebra, Spectral Decomposition, Operator Functions.
Recall that $\sigma_1\equiv X, \sigma_2\equiv Y$, and $\sigma_3\equiv Z$.

First, we compute $\mathbf{v} \cdot \boldsymbol{\sigma}$ in matrix form:

$$\mathbf{v} \cdot \boldsymbol{\sigma} = v_1\sigma_1 + v_2\sigma_2 + v_3\sigma_3 = v_1 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + v_2 \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + v_3 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} v_3 & v_1 - iv_2 \\ v_1 + iv_2 & -v_3 \end{bmatrix}$$

In order to compute the complex exponential of this matrix, we will want to find its spectral decomposition. Using the characterstic equation to find the eigenvalues, we have:

$$\begin{aligned}
\det(\mathbf{v} \cdot \boldsymbol{\sigma} - I\lambda) = 0 &\implies \det \begin{bmatrix} v_3 - \lambda & v_1 - iv_2 \\ v_1 + iv_2 & -v_3 - \lambda \end{bmatrix} = 0 \\
&\implies (v_3 - \lambda)(-v_3 - \lambda) - (v_1 - iv_2)(v_1 + iv_2) = 0 \\
&\implies \lambda^2 - v_3^2 - v_1^2 - v_2^2 = \lambda^2 - (v_1^2 + v_2^2 + v_3^2) = 0 \\
&\implies \lambda^2 - 1 = 0 \\
&\implies \lambda_1 = 1, \lambda_2 = -1
\end{aligned}$$

where in the second-to-last implication we use the fact that $\mathbf{v}$ is a unit vector. Letting $|v_1\rangle, |v_2\rangle$ be the associated eigenvectors, $\mathbf{v} \cdot \boldsymbol{\sigma}$ has spectral decomposition:

$$\mathbf{v} \cdot \boldsymbol{\sigma} = |v_1\rangle\langle v_1| - |v_2\rangle\langle v_2|$$

Applying the complex exponentiation operator, we then have that:

$$\exp(i\theta\mathbf{v} \cdot \boldsymbol{\sigma}) = \exp(i\theta) |v_1\rangle\langle v_1| + \exp(-i\theta) |v_2\rangle\langle v_2| .$$

Using Euler's formula, we then have that:

$$\begin{aligned}
\exp(i\theta\mathbf{v} \cdot \boldsymbol{\sigma}) &= (\cos\theta + i\sin\theta) |v_1\rangle\langle v_1| + (\cos\theta - i\sin\theta) |v_2\rangle\langle v_2| \\
&= \cos(\theta) \left(|v_1\rangle\langle v_1| + |v_2\rangle\langle v_2| \right) + i\sin(\theta) \left(|v_1\rangle\langle v_1| - |v_2\rangle\langle v_2| \right) \\
&= \cos(\theta)I + i\sin(\theta)\mathbf{v} \cdot \boldsymbol{\sigma}.
\end{aligned}$$

Where in the last line we use the completeness relation and the spectral decomposition. □

Exercise 2.36

Show that the Pauli matrices except for I have trace zero.

Solution

Concepts Involved: Linear Algebra, Trace.

We have that:

$$\operatorname{tr}(X) = \operatorname{tr}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = 0 + 0 = 0$$
$$\operatorname{tr}(Y) = \operatorname{tr}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = 0 + 0 = 0$$
$$\operatorname{tr}(Z) = \operatorname{tr}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = 1 - 1 = 0$$

□

Exercise 2.37: Cyclic property of the trace

If A and B are two linear operators show that

$$\operatorname{tr}(AB) = \operatorname{tr}(BA)$$

Solution

Concepts Involved: Linear Algebra, Trace.

Let A, B be linear operators. Then, $C = AB$ has matrix representation with entries $C_{ij} = \sum_k A_{ik}B_{kj}$ and $D = BA$ has matrix representation with entries $D_{ij} = \sum_k B_{ik}A_{kj}$. We then have that:

$$\operatorname{tr}(AB) = \operatorname{tr}(C) = \sum_i C_{ii} = \sum_i \sum_k A_{ik}B_{ki} = \sum_k \sum_i B_{ki}A_{ik} = \sum_k D_{kk} = \operatorname{tr}(D) = \operatorname{tr}(BA)$$

□

Exercise 2.38: Linearity of the trace

If A and B are two linear operators, show that

$$\operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B)$$

and if z is an arbitrary complex number show that

$$\operatorname{tr}(zA) = z\operatorname{tr}(A).$$

Solution

Concepts Involved: Linear Algebra, Trace.

From the definition of trace, we have that:

$$\operatorname{tr}(A+B) = \sum_i (A+B)_{ii} = \sum_i A_{ii} + B_{ii} = \sum_i A_{ii} + \sum_i B_{ii} = \operatorname{tr}(A) + \operatorname{tr}(B)$$

$$\operatorname{tr}(zA) = \sum_i (zA)_{ii} = z\sum_i A_{ii} = z\operatorname{tr}(A)$$

□

Exercise 2.39: The Hilbert-Schmidt inner product on operators

The set L_V of linear operators on Hilbert space V is obviously a vector space - the sum of two linear operators is a linear operator, zA is a linear operator if A is a linear operator and z is a complex number, and there is a zero element 0. An important additional result is that the vector space L_V can be given a natural inner product structure, turning it into a Hilbert space.

(1) Show that the function $(\cdot,\cdot)$ on $L_V \times L_V$ defined by

$$(A, B) \equiv \operatorname{tr}\left(A^\dagger B\right)$$

is an inner product function. This inner product is known as the *Hilbert-Schmidt* or *trace* inner product.

(2) If V has d dimensions show that L_V has dimension d^2.

(3) Find an orthonormal basis of Hermitian matrices for the Hilbert space L_V.

Solution

Concepts Involved: Linear Algebra, Trace, Inner Products, Hermitian Operators, Bases

(1) We show that $(\cdot,\cdot)$ satisfies the three properties of an inner product. Showing that it is linear in the second argument, we have that:

$$\left(A, \sum_i \lambda_i B_i\right) = \operatorname{tr}\left(A\sum_i \lambda_i B_i\right) = \sum_i \lambda_i \operatorname{tr}(AB_i) = \sum_i \lambda_i (A, B_i)$$

where in the second to last equality we use the result of Exercise 2.38. To see that it is conjugate-symmetric, we have that:

$$(A, B) = \operatorname{tr}\left(A^\dagger B\right) = \operatorname{tr}\left((B^\dagger A)^\dagger\right) = \operatorname{tr}\left(B^\dagger A\right)^* = (B, A)^*$$

Finally, to show positive definitness, we have that:

$$(A, A) = \operatorname{tr}\left(A^\dagger A\right) = \sum_i\sum_k A^\dagger_{ik}A_{ki} = \sum_i\sum_k A^*_{ki}A_{ki} = \sum_i\sum_k |A_{ki}|^2 \geq 0$$

so we conclude that $(\cdot,\cdot)$ is an inner product function.

(2) Suppose V has d dimensions. Then, the elements of L_V which consist of linear operators $A : V \mapsto V$ have representations as $d \times d$ matrices. There are d^2 such linearly independent matrices (take the matrices with 1 in one of the d^2 entries and 0 elsewhere), and we conclude that L_V has d^2 linearly

independent vectors and hence dimension d^2.

(3) As discussed in the previous part of the question, one possible basis for this vector space would be $|v_i\rangle\langle v_j|$ where $|v_k\rangle$ form an orthonormal basis of V with $i, j \in \{1, \dots d\}$. These of course are just matrices with 1 in one entry and 0 elsewhere. It is easy to see that this is a basis as for any $A \in L_V$ we can write $A = \sum_{ij} \lambda_{ij} |v_i\rangle\langle v_j|$. We can verify that these are orthonormal; suppose $|v_{i_1}\rangle\langle v_{j_1}| \neq |v_{i_2}\rangle\langle v_{j_2}|$. Then, we have that:

$$\begin{aligned}\big(|v_{i_1}\rangle\langle v_{j_1}|, |v_{i_2}\rangle\langle v_{j_2}|\big) &= \operatorname{tr}\Big((|v_{i_1}\rangle\langle v_{j_1}|)^\dagger |v_{i_2}\rangle\langle v_{j_2}|\Big) \\ &= \operatorname{tr}\big(|v_{j_1}\rangle\langle v_{i_1}|v_{i_2}\rangle\langle v_{j_2}|\big)\end{aligned}$$

If $|v_{i_1}\rangle \neq |v_{i_2}\rangle$, then the above expression reduces to $\operatorname{tr}(0) = 0$. If $|v_{i_1}\rangle = |v_{i_2}\rangle$, then it follows that $|v_{j_1}\rangle \neq |v_{j_2}\rangle$ (else this would contradict $|v_{i_1}\rangle\langle v_{j_1}| \neq |v_{i_2}\rangle\langle v_{j_2}|$) and in this case we have that:

$$\begin{aligned}\big(|v_{i_1}\rangle\langle v_{j_1}|, |v_{i_2}\rangle\langle v_{j_2}|\big) &= \operatorname{tr}\big(|v_{j_1}\rangle\langle v_{i_1}|v_{i_2}\rangle\langle v_{j_2}|\big) \\ &= \operatorname{tr}\big(|v_{j_1}\rangle\langle v_{j_2}|\big) \\ &= 0\end{aligned}$$

So we therefore have that the inner product of two non-identical elements in the basis is zero. Furthermore, we have that:

$$\big(|v_{i_1}\rangle\langle v_{j_1}|, |v_{i_1}\rangle\langle v_{j_1}|\big) = \operatorname{tr}\big(|v_{i_1}\rangle\langle v_{j_1}|\ |v_{i_1}\rangle\langle v_{j_1}|\big) = \operatorname{tr}\big(|v_{i_1}\rangle\langle v_{i_1}|\big) = 1$$

so we confirm that this basis is orthonormal. However, evidently this basis is *not* Hermitian as if $i \neq j$, then $(|v_i\rangle\langle v_j|)^\dagger = |v_j\rangle\langle v_i| \neq |v_i\rangle\langle v_j|$. To fix this, we can modify our basis slightly. We keep the diagonal entries as is (as these are indeed Hermitian!) but for the off-diagonals, we replace every pair of basis vectors $|v_i\rangle\langle v_j|$, $|v_j\rangle\langle v_i|$ with:

$$\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}}, \quad i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}.$$

A quick verification shows that these are indeed Hermitian:

$$\begin{aligned}\left(\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}}\right)^\dagger &= \frac{(|v_i\rangle\langle v_j|)^\dagger + (|v_j\rangle\langle v_i|)^\dagger}{\sqrt{2}} = \frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}} \\ \left(i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\right)^\dagger &= -i\frac{(|v_i\rangle\langle v_j|)^\dagger - (|v_j\rangle\langle v_i|)^\dagger}{\sqrt{2}} = i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\end{aligned}$$

It now suffices to show that these new vectors (plus the diagonals) form a basis and are orthonormal. To see that these form a basis, observe that:

$$\begin{aligned}\frac{1}{\sqrt{2}}\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}} - \frac{i}{\sqrt{2}}\left(i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\right) &= |v_i\rangle\langle v_j| \\ \frac{1}{\sqrt{2}}\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}} + \frac{i}{\sqrt{2}}\left(i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\right) &= |v_j\rangle\langle v_i|\end{aligned}$$

and since we know that $|v_i\rangle\langle v_j|$ for all $i, j \in \{1, \dots d\}$ form a basis, this newly defined set of vectors must be a basis as well. Furthermore, since the new basis vectors are constructed from orthogonal $|v_i\rangle\langle v_j|$, the newly defined vectors will be orthogonal to each other if $i_1, j_1 \neq i_2, j_2$. The only things left to check is that for any choice of i, j that:

$$\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}} \text{ and } i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}.$$

are orthogonal, and that these vectors are normalized. Checking the orthogonality, we have:

$$\begin{aligned}
\left(\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}}, i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\right) &= \operatorname{tr}\left(\left(\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}}\right)\left(i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\right)\right) \\
&= \frac{i}{\sqrt{2}} \operatorname{tr}\big(|v_j\rangle\langle v_j| - |v_i\rangle\langle v_i|\big) \\
&= 0.
\end{aligned}$$

And checking the normalization, we have that:

$$\begin{aligned}
\left(\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}}, \frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}}\right) &= \operatorname{tr}\left(\left(\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}}\right)\left(\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}}\right)\right) \\
&= \frac{1}{2} \operatorname{tr}\big(|v_i\rangle\langle v_i| + |v_j\rangle\langle v_j|\big) \\
&= 1
\end{aligned}$$

$$\begin{aligned}
\left(i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}, i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\right) &= \operatorname{tr}\left(\left(i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\right)\left(i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\right)\right) \\
&= -\frac{1}{2} \operatorname{tr}\big(-|v_i\rangle\langle v_i| - |v_j\rangle\langle v_j|\big) \\
&= 1
\end{aligned}$$

□

Exercise 2.40: Commutation relations for the Pauli matrices

Verify the commutation relations

$$[X, Y] = 2iZ; \quad [Y, Z] = 2iX; \quad [Z, X] = 2iY$$

There is an elegant way of writing this using ϵ_{jkl}, the antisymmetric tensor on three indices, for which $\epsilon_{jkl} = 0$ except for $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$, and $\epsilon_{321} = \epsilon_{213} = \epsilon_{132} = -1$:

$$[\sigma_j, \sigma_k] = 2i \sum_{l=1}^{3} \epsilon_{jkl} \sigma_l$$

Solution

Concepts Involved: Linear Algebra, Commutators.

We verify the proposed relations via computation in the computational basis:

$$[X,Y] = XY - YX = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} - \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} - \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} = 2iZ$$

$$[Y,Z] = YZ - ZY = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} - \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix} = 2iX$$

$$[Z,X] = ZX - XZ = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = 2iY$$

□

Exercise 2.41: Anti-commutation relations for the Pauli matrices

Verify the anticommutation relations

$$\{\sigma_i, \sigma_j\} = 0$$

Where $i \neq j$ are both chosen from the set $1, 2, 3$. Also verify that $(i = 0, 1, 2, 3)$

$$\sigma_i^2 = I$$

Solution

Concepts Involved: Linear Algebra, Anticommutators.

We again verify the proposed relations via computation in the computational basis:

$$\{X,Y\} = XY + YX = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} + \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\{Y,Z\} = YZ + ZY = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} + \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\{Z,X\} = ZX + XZ = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

This proves the first claim as $\{A,B\} = AB + BA = BA + AB = \{B,A\}$ and the other 3 relations are equivalent to the ones already proven. Verifying the second claim, we have:

$$I^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$X^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$Y^2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$Z^2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

□

Remark: Note that we can write this result consicely as $\{\sigma_j, \sigma_k\} = 2\delta_{ij} I$

Exercise 2.42

Verify that

$$AB = \frac{[A,B] + \{A,B\}}{2}$$

Solution

Concepts Involved: Linear Algebra, Commutators, Anticommutators.

By algebraic manipulation we obtain:

$$AB = \frac{AB + AB}{2} + \frac{BA - BA}{2} = \frac{(AB - BA) + (AB + BA)}{2} = \frac{[A,B] + \{A,B\}}{2}$$

□

Exercise 2.43

Show that for $j, k = 1, 2, 3$,

$$\sigma_j \sigma_k = \delta_{jk} I + i \sum_{l=1}^{3} \epsilon_{jkl} \sigma_l.$$

Solution

Concepts Involved: Linear Algebra, Commutators, Anticommutators.

Applying the results of Exercises 2.40, 2.41, and 2.42, we have:

$$\begin{aligned}\sigma_j\sigma_k &= \frac{[\sigma_j,\sigma_k] + \{\sigma_j,\sigma_k\}}{2} \\ &= \frac{2i\sum_{l=1}^{3}\epsilon_{jkl}\sigma_l + 2\delta_{ij}I}{2} \\ &= \delta_{ij}I + i\sum_{l=1}^{3}\epsilon_{jkl}\sigma_l\end{aligned}$$

□

Exercise 2.44

Suppose $[A,B]=0, \{A,B\}=0$, and A is invertible. Show that B must be 0.

Solution

Concepts Involved: Linear Algebra, Commutators, Anticommutators.

By assumption, we have that:

$$\begin{aligned}[A,B] &= AB - BA = 0 \\ \{A,B\} &= AB + BA = 0.\end{aligned}$$

Adding the first line to the second we have:

$$2AB = 0 \implies AB = 0.$$

A^{-1} exists by the invertibility of A, so multiplying by A^{-1} on the left we have:

$$A^{-1}AB = A^{-1}0 \implies IB = 0 \implies B = 0.$$

□

Exercise 2.45

Show that $[A,B]^\dagger = [B^\dagger, A^\dagger]$.

Solution

Concepts Involved: Linear Algebra, Commutators, Adjoints.

Using the properties of the adjoint, we have:

$$[A,B]^\dagger = (AB-BA)^\dagger = (AB)^\dagger - (BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger = [B^\dagger, A^\dagger]$$

□

Exercise 2.46

Show that $[A, B] = -[B, A]$.

Solution

Concepts Involved: Linear Algebra, Commutators

By the definition of the commutator:

$$[A, B] = AB - BA = -(BA - AB) = -[B, A]$$

$\square$

Exercise 2.47

Suppose A and B are Hermitian. Show that $i[A, B]$ is Hermitian.

Solution

Concepts Involved: Linear Algebra, Commutators, Hermitian Operators

Suppose A, B are Hermitian. Using the results of Exercises 2.45 and 2.46, we have:

$$(i[A, B])^\dagger = -i([A, B])^\dagger = -i[B^\dagger, A^\dagger] = i[A^\dagger, B^\dagger] = i[A, B].$$

$\square$

Exercise 2.48

What is the polar decomposition of a positive matrix P? Of a unitary matrix U? Of a Hermitian matrix, H?

Solution

Concepts Involved: Linear Algebra, Polar Decomposition, Positive Operators, Unitary Operators, Hermitian Operators

If P is a positive matrix, then no calculation is required; $P = IP = PI$ is the polar decomposition (as I is unitary and P is positive). If U is a unitary matrix, then $J = \sqrt{U^\dagger U} = \sqrt{I} = I$ and $K = \sqrt{UU^\dagger} = \sqrt{I} = I$ so the polar decomposition is $U = UI = IU$ (where U is unitary and I is positive). If H is hermitian, we then have that:

$$J = \sqrt{H^\dagger H} = \sqrt{H^2} = \sqrt{\sum_i \lambda_i^2 |i\rangle\langle i|} = \sum_i |\lambda_i| |i\rangle\langle i|$$

and $K = \sqrt{HH^\dagger} = \sum_i |\lambda_i| |i\rangle\langle i|$ in the same way. Hence the polar decomposition is $H = U \sum_i |\lambda_i| |i\rangle\langle i| = \sum_i |\lambda_i| |i\rangle\langle i| U$. $\square$

Exercise 2.49

Express the polar decomposition of a normal matrix in the outer product representation.

Solution

Concepts Involved: Linear Algebra, Polar Decomposition, Outer Products

Let A be a normal matrix. Then, A has spectral decomposition $A = \sum_i \lambda_i |i\rangle\langle i|$. Therefore, we have that:

$$A^\dagger A = AA^\dagger = \sum_i \sum_{i'} \lambda_i \lambda_{i'}^* |i\rangle\langle i| |i'\rangle\langle i'| = \sum_i \sum_{i'} \lambda_i \lambda_{i'}^* |i\rangle\langle i'| \delta_{ii'} = \sum_i |\lambda_i|^2 |i\rangle\langle i|$$

We then have that:

$$J = \sqrt{A^\dagger A} = \sqrt{\sum_i |\lambda_i|^2 |i\rangle\langle i|} = \sum_i |\lambda_i| |i\rangle\langle i|$$

and $K = \sum_i |\lambda_i| |i\rangle\langle i|$ identically. Furthermore, U is unitary, so it also has a spectral decomposition of $\sum_j \mu_j |j\rangle\langle j|$. Hence we have the polar decomposition in the outer product representation as:

$$\begin{aligned}
A &= UJ = KU \\
A &= U \sum_i |\lambda_i| |i\rangle\langle i| \sum_j = \sum_i |\lambda_i| |i\rangle\langle i| \sum_j U \\
A &= \sum_j \sum_i \mu_j |\lambda_i| |j\rangle\langle j|i\rangle\langle i| = \sum_i \sum_j |\lambda_i| \mu_j |i\rangle\langle i|j\rangle\langle j|
\end{aligned}$$

□

Exercise 2.50

Find the left and right polar decompositions of the matrix

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

Solution

Concepts Involved: Linear Algebra, Polar Decomposition.

Let $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. We start with the left polar decomposition, and hence find $J = \sqrt{A^\dagger A}$. In order to do this, we find the spectral decompositions of $A^\dagger A$ and $AA^\dagger$.

$$\det\Big(A^\dagger A - I\lambda\Big) = 0 \implies \det\left(\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix} - \begin{bmatrix}\lambda & 0\\ 0 & \lambda\end{bmatrix}\right) = 0 \implies \det\begin{bmatrix}2-\lambda & 1\\ 1 & 1-\lambda\end{bmatrix} = 0$$
$$\implies \lambda^2 - 3\lambda + 1 = 0$$
$$\implies \lambda_1 = \frac{3+\sqrt{5}}{2}, \lambda_2 = \frac{3-\sqrt{5}}{2}$$

Solving for the eigenvectors, we have:

$$\begin{bmatrix}2 - \frac{3+\sqrt{5}}{2} & 1\\ 1 & 1 - \frac{3+\sqrt{5}}{2}\end{bmatrix}|v_1\rangle = \mathbf{0} \implies |v_1\rangle = \begin{bmatrix}1+\sqrt{5}\\ 2\end{bmatrix}$$
$$\begin{bmatrix}2 - \frac{3-\sqrt{5}}{2} & 1\\ 1 & 1 - \frac{3-\sqrt{5}}{2}\end{bmatrix}|v_2\rangle = \mathbf{0} \implies |v_2\rangle = \begin{bmatrix}1-\sqrt{5}\\ 2\end{bmatrix}$$

Normalizing, we get:

$$|v_1\rangle = \frac{1}{\sqrt{10+2\sqrt{5}}}\begin{bmatrix}1+\sqrt{5}\\ 2\end{bmatrix}, \quad |v_2\rangle = \frac{1}{\sqrt{10-2\sqrt{5}}}\begin{bmatrix}1-\sqrt{5}\\ 2\end{bmatrix}$$

The spectral decomposition of $A^\dagger A$ is therefore:

$$A^\dagger A = \lambda_1 |v_1\rangle\langle v_1| + \lambda_2 |v_2\rangle\langle v_2|$$

Calculating J, we therefore have:

$$J = \sqrt{A^\dagger A} = \sqrt{\lambda_1}\,|v_1\rangle\langle v_1| + \sqrt{\lambda_2}\,|v_2\rangle\langle v_2| = \frac{1}{\sqrt{5}}\begin{bmatrix}3 & 1\\ 1 & 2\end{bmatrix}$$

The last equality is not completely trivial, but the algebra is tedious so we invite the reader to use a symbolic calculator, as we have. We make the observation that:

$$A = UJ \implies U = AJ^{-1}$$

So calculating J^{-1}, we have:

$$J^{-1} = \frac{1}{\sqrt{\lambda_1}}|v_1\rangle\langle v_1| + \frac{1}{\sqrt{\lambda_2}}|v_2\rangle\langle v_2| = \frac{1}{\sqrt{5}}\begin{bmatrix}2 & -1\\ -1 & 3\end{bmatrix}$$

Where we again have used the help of a symbolic calculator. Calculating U, we then have that:

$$U = AJ^{-1} = \begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}\frac{1}{\sqrt{5}}\begin{bmatrix}2 & -1\\ -1 & 3\end{bmatrix} = \frac{1}{\sqrt{5}}\begin{bmatrix}2 & -1\\ 1 & 2\end{bmatrix}$$

Hence the left polar decomposition of A is given by:

$$A = UJ = \left(\frac{1}{\sqrt{5}}\begin{bmatrix}2 & -1\\ 1 & 2\end{bmatrix}\right)\left(\frac{1}{\sqrt{5}}\begin{bmatrix}3 & 1\\ 1 & 2\end{bmatrix}\right)$$

We next solve for the right polar decomposition. We could repeat the procedure of solving for the spectral decomposition of $AA^\dagger$, but we take a shortcut; since we know the K that satisfies:

$$A = KU$$

will be unique, and U is unitary, we can simply multiply both sides of the above equation on the right by $U^{-1} = U^\dagger$ to obtain K. Hence:

$$K = AU^\dagger = \begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}\frac{1}{\sqrt{5}}\begin{bmatrix}2 & 1\\ -1 & 2\end{bmatrix} = \frac{1}{\sqrt{5}}\begin{bmatrix}2 & 1\\ 1 & 3\end{bmatrix}.$$

Therefore the right polar decomposition of A is given by:

$$A = KU = \left(\frac{1}{\sqrt{5}}\begin{bmatrix}2 & 1\\ 1 & 3\end{bmatrix}\right)\left(\frac{1}{\sqrt{5}}\begin{bmatrix}2 & -1\\ 1 & 2\end{bmatrix}\right)$$

□

Exercise 2.51

Verify that the Hadamard gate H is unitary.

Solution

Concepts Involved: Linear Algebra, Unitary Operators

We observe that:

$$H^\dagger H = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

showing that H is indeed unitary. □

Remark: The above calculation shows that H is also Hermitian and Idempotent.

Exercise 2.52

Verify that $H^2 = I$.

Solution

Concepts Involved: Linear Algebra

See the calculation and remark in the previous exercise. $\square$

Exercise 2.53

What are the eigenvalues and eigenvectors of H?

Solution

Concepts Involved: Linear Algebra, Eigenvalues, Eigenvectors

Using the characteristic equation to find the eigenvalues, we have:

$$\begin{aligned}\det(H - I\lambda) = 0 \implies \det\begin{bmatrix}\frac{1}{\sqrt{2}} - \lambda & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} - \lambda\end{bmatrix} = 0 &\implies \lambda^2 - 1 = 0 \\ &\implies \lambda_1 = 1, \lambda_2 = -1\end{aligned}$$

Finding the eigenvectors, we then have:

$$\begin{bmatrix}\frac{1}{\sqrt{2}} - 1 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} - 1\end{bmatrix}|v_1\rangle = \mathbf{0} \implies |v_1\rangle = \begin{bmatrix}1 + \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}}\end{bmatrix}$$
$$\begin{bmatrix}\frac{1}{\sqrt{2}} + 1 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} + 1\end{bmatrix}|v_2\rangle = \mathbf{0} \implies |v_2\rangle = \begin{bmatrix}-1 + \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}}\end{bmatrix}$$

Normalizing, we have:

$$|v_1\rangle = \frac{1}{\sqrt{2+\sqrt{2}}}\begin{bmatrix}1 + \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}}\end{bmatrix}, |v_2\rangle = \frac{1}{\sqrt{2-\sqrt{2}}}\begin{bmatrix}-1 + \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}}\end{bmatrix}$$

$\square$

Exercise 2.54

Suppose A and B are commuting Hermitian operators. Prove that $\exp(A)\exp(B) = \exp(A+B)$. (*Hint:* Use the results of Section 2.1.9.)

Solution

Concepts Involved: Linear Algebra, Operator Functions, Simultaneous Diagonalization

Since A, B commute, they can be simulatneously diagonalized; that is, there exists some orthonormal basis $|i\rangle$ of V such that $A = \sum_i a_i |i\rangle\langle i|$ and $B = \sum_i b_i |i\rangle\langle i|$. Hence, using the definition of operator functions, we have that:

$$
\begin{aligned}
\exp(A)\exp(B) &= \exp\left(\sum_i a_i \, |i\rangle\langle i|\right)\exp\left(\sum_{i'} b_{i'} \, |i'\rangle\langle i'|\right) \\
&= \sum_i \sum_{i'} \exp(a_i)\exp(b_{i'})|i\rangle\langle i|i'\rangle\langle i'| \\
&= \sum_i \sum_{i'} \exp(a_i)\exp(b_{i'})|i\rangle\langle i'|\delta_{ii'} \\
&= \sum_i \exp(a_i)\exp(b_i)\,|i\rangle\langle i| \\
&= \sum_i \exp(a_i+b_i)\,|i\rangle\langle i| \\
&= \exp\left(\sum_i (a_i+b_i)\,|i\rangle\langle i|\right) \\
&= \exp(A+B)
\end{aligned}
$$

□

Exercise 2.55

Prove that $U(t_1, t_2)$ defined in Equation (2.91) is unitary.

Solution

Concepts Involved: Linear Algebra, Unitary Operators, Spectral Decomposition, Operator Functions.

Since the Hamiltonian H is Hermitian, it is normal and hence has spectral decomposition:

$$
H = \sum_E E\,|E\rangle\langle E|
$$

where all E are real by the Hermicity of H, and $|E\rangle$ is an orthonormal basis of the Hilbert space. We then have that:

$$
\begin{aligned}
U(t_1,t_2) \equiv \exp\left[\frac{-iH(t_2-t_1)}{\hbar}\right] &= \exp\left[\frac{-i\sum_E E\,|E\rangle\langle E|\,(t_2-t_1)}{\hbar}\right] \\
&= \sum_E \exp\left(\frac{-iE(t_2-t_1)}{\hbar}\right)|E\rangle\langle E|
\end{aligned}
$$

Hence calculating $U^\dagger(t_1, t_2)$ we have:

$$\begin{aligned}U^\dagger(t_1,t_2) = \left(\sum_E \exp\left(\frac{-iE(t_2-t_1)}{\hbar}\right)|E\rangle\langle E|\right)^\dagger &= \sum_E \left(\exp\left(\frac{-iE(t_2-t_1)}{\hbar}\right)\right)^* \left(|E\rangle\langle E|\right)^\dagger \\ &= \sum_E \exp\left(\frac{iE(t_2-t_1)}{\hbar}\right)|E\rangle\langle E|\end{aligned}$$

Therefore computing $U^\dagger(t_1, t_2)U(t_2, t_1)$ we have:

$$\begin{aligned}U^\dagger(t_2,t_1)U(t_2,t_1) &= \left(\sum_E \exp\left(\frac{-iE(t_2-t_1)}{\hbar}\right)|E\rangle\langle E|\right)\left(\sum_{E'} \exp\left(\frac{iE'(t_2-t_1)}{\hbar}\right)|E'\rangle\langle E'|\right) \\ &= \sum_E\sum_{E'} \exp\left(\frac{-iE(t_2-t_1)}{\hbar}\right)\exp\left(\frac{iE'(t_2-t_1)}{\hbar}\right)\delta_{EE'}|E\rangle\langle E'| \\ &= \sum_E \exp\left(\frac{-iE(t_2-t_1)}{\hbar}\right)\exp\left(\frac{iE(t_2-t_1)}{\hbar}\right)|E\rangle\langle E| \\ &= \sum_E |E\rangle\langle E| \\ &= I\end{aligned}$$

where in the second equality we use the fact that the eigenstates are orthogonal. We conclude that U is unitary. □

Exercise 2.56

Use the spectral decomposition to show that $K \equiv -i\log(U)$ is Hermitian for any unitary U, and thus $U = \exp(iK)$ for some Hermitian K.

Solution

Concepts Involved: Linear Algebra, Hermitian Operators, Unitary Operators, Spectral Decomposition, Operator Functions.

Suppose U is unitary. Then, U is normal and hence has spectral decomposition:

$$U = \sum_j \lambda_j |j\rangle\langle j|$$

where $|j\rangle$ are eigenvectors of U with eigenvalues λ_j, and $|j\rangle$ forms an orthonormal basis of the Hilbert space. By Exercise 2.18, all eigenvalues of unitary operators have eigenvalues of modulus 1, so we can let $\lambda_j = \exp(i\theta_j)$ where $\theta_j \in \mathbb{R}$ and hence write the above as:

$$U = \sum_j \exp(i\theta_j)|j\rangle\langle j|$$

We then have that:

$$K \equiv -i\log(U) = -i\log\left(\sum_j \exp(i\theta_j)\,|j\rangle\!\langle j|\right) = \sum_j -i\log\Big(\exp(i\theta_j)\Big)\,|j\rangle\!\langle j| = \sum_j -i(i\theta_j)\,|j\rangle\!\langle j|$$
$$= \sum_j \theta_j\,|j\rangle\!\langle j|$$

We then observe that:

$$K^\dagger = \left(\sum_j \theta_j\,|j\rangle\!\langle j|\right)^\dagger = \sum_j \theta_j\,|j\rangle\!\langle j|$$

as the θ_js are real and $(|j\rangle\langle j|)^\dagger = |j\rangle\!\langle j|$. Hence K is Hermitian. Then, multiplying both sides in $K = -i\log(U)$ by i and exponentiating both sides, we obtain the desired relation. □

Exercise 2.57: Cascaded measurements are single measurements

Suppose $\{L_l\}$ and $\{M_m\}$ are two sets of measurement operators. Show that a measurement defined by the measurement operators $\{L_l\}$ followed by a measurement defined by the measurement operators $\{M_m\}$ is physically equivalent to a single measurement defined by measurement operators $\{N_{lm}\}$ with the representation $N_{lm} \equiv M_m L_l$.

Solution

Concepts Involved: Linear Algebra, Quantum Measurement.

Suppose we have (normalized) initial quantum state $|\psi_0\rangle$. Then, the state after measurement of L_l is given by definition to be:

$$|\psi_0\rangle \mapsto |\psi_1\rangle = \frac{L_l|\psi_0\rangle}{\sqrt{\langle\psi_0|L_l^\dagger L_l|\psi_0\rangle}}.$$

The state after measurement of M_m on $|\psi_1\rangle$ is then given to be:

$$|\psi_1\rangle \mapsto |\psi_2\rangle = \frac{M_m|\psi_1\rangle}{\sqrt{\langle\psi_1|M_m^\dagger M_m|\psi_1\rangle}} = \frac{M_m\left(\frac{L_l|\psi_0\rangle}{\sqrt{\langle\psi_0|L_l^\dagger L_l|\psi_0\rangle}}\right)}{\sqrt{\left(\frac{L_l^\dagger\langle\psi_0|}{\sqrt{\langle\psi_0|L_l^\dagger L_l|\psi_0\rangle}}\right)M_m^\dagger M_m\left(\frac{L_l|\psi_0\rangle}{\sqrt{\langle\psi_0|L_l^\dagger L_l|\psi_0\rangle}}\right)}}$$
$$= \frac{M_m L_l|\psi_0\rangle}{\sqrt{\langle\psi_0|L_l^\dagger L_l|\psi_0\rangle}}\frac{\sqrt{\langle\psi_0|L_l^\dagger L_l|\psi_0\rangle}}{\sqrt{\langle\psi_0|L_l^\dagger M_m^\dagger M_m L_l^\dagger|\psi_0\rangle}}$$
$$= \frac{M_m L_l|\psi_0\rangle}{\sqrt{\langle\psi_0|L_l^\dagger M_m^\dagger M_m L_l^\dagger|\psi_0\rangle}}.$$

Conversely, the state of $|\psi_0\rangle$ after measurement of $N_{lm} = M_m L_l$ is given by:

$$|\psi_0\rangle \mapsto |\psi_3\rangle = \frac{M_m L_l|\psi_0\rangle}{\sqrt{\langle\psi_0|L_l^\dagger M_m^\dagger M_m L_l|\psi_0\rangle}}.$$

We see that $|\psi_2\rangle = |\psi_3\rangle$ (that is, the cascaded measurment produces the same result as the single measurement), proving the claim. □

Exercise 2.58

Suppose we prepare a quantum system in an eigenstate $|\psi\rangle$ of some observable M with corresponding eigenvalue m. What is the average observed value of M, and the standard deviation?

Solution

Concepts Involved: Linear Algebra, Quantum Measurement, Expectation, Standard Deviation.

By the definition of expectation, we have that:

$$\langle M\rangle_{|\psi\rangle} = \langle\psi|M|\psi\rangle = \langle\psi|m|\psi\rangle = m\langle\psi|\psi\rangle = m$$

Where in the second equality we use that $|\psi\rangle$ is an eigenstate of M with eigenvalue m, and in the last equality we use that $|\psi\rangle$ is a normalized quantum state. Next, calculating $\langle M^2\rangle_{|\psi\rangle}$, we have:

$$\left\langle M^2\right\rangle_{|\psi\rangle} = \langle\psi|M^2|\psi\rangle = \langle\psi|MM|\psi\rangle = \langle\psi|M^\dagger M|\psi\rangle = \langle\psi|m^* m|\psi\rangle = \langle\psi|m^2|\psi\rangle = m^2\langle\psi|\psi\rangle = m^2.$$

Note that we have used the fact that M is Hermitian (it is an observable) to use that $M^\dagger = M$ and $m^* = m$ as all eigenvalues of Hermitian operators are real. Now calculating the standard deviation, we have:

$$\Delta(M) = \sqrt{\langle M^2\rangle - \langle M\rangle^2} = \sqrt{m^2 - (m)^2} = 0$$

□

Exercise 2.59

Suppose we have qubit in the state $|0\rangle$, and we measure the observable X. What is the average value of X? What is the standard deviation of X?

Solution

Concepts Involved: Linear Algebra, Quantum Measurement, Projective Measurement, Expectation, Standard Deviation.

By the definition of expectation, we have:

$$\langle X\rangle_{|0\rangle} = \langle 0|X|0\rangle = \langle 0|1\rangle = 0$$

Next calculating $\langle X^2 \rangle_{|0\rangle}$, we have:

$$\left\langle X^2 \right\rangle_{|0\rangle} = \langle 0|XX|0\rangle = \langle 1|1\rangle = 1$$

Hence the standard deviation of X is given by:

$$\Delta(X) = \sqrt{\langle X^2\rangle - \langle X\rangle^2} = \sqrt{1-0} = 1$$

□

Exercise 2.60

Show that $\mathbf{v}\cdot\boldsymbol{\sigma}$ has eigenvalues ± 1, and that the projectors into the corresponding eigenspaces are given by $P_\pm = (I \pm \mathbf{v}\cdot\boldsymbol{\sigma})/2$.

Solution

Concepts Involved: Eigenvalues, Projectors.

Let $|v\rangle$ be a unit vector. We already showed in Exercise 2.35 that $\mathbf{v}\cdot\boldsymbol{\sigma}$ has eigenvalues $\lambda_+ = 1, \lambda_- = -1$. We next prove a general statement; namely, that for a observable on a 2-dimensional Hilbert space with eigenvalues $\lambda_\pm = \pm 1$ has projectors

$$P_\pm = \frac{I \pm O}{2}$$

To see this is the case, let $P_+ = |o_+\rangle\langle o_+|$, $P_- = |o_-\rangle\langle o_-|$, $I = |o_+\rangle\langle o_+| + |o_-\rangle\langle o_-|$, and $O = |o_+\rangle\langle o_+| - |o_-\rangle\langle o_-|$. We then have that:

$$\frac{I+O}{2} = \frac{|o_+\rangle\langle o_+| - |o_-\rangle\langle o_-|}{2} = |o_+\rangle\langle o_+| = P_+$$
$$\frac{I-O}{2} = \frac{|o_-\rangle\langle o_-| + |o_-\rangle\langle o_-| - |o_+\rangle\langle o_+| + |o_-\rangle\langle o_-|}{2} = |o_-\rangle\langle o_-| = P_-$$

Hence the general statement is proven. Applying this to $O = \mathbf{v}\cdot\boldsymbol{\sigma}$ (which is indeed Hermitian and hence an observable as each of X, Y, Z are Hermitian), we get that:

$$P_\pm = \frac{I \pm \mathbf{v}\cdot\boldsymbol{\sigma}}{2}$$

as claimed. □

Exercise 2.61

Calculate the probability of obtaining the result $+1$ for a measurement of $\mathbf{v}\cdot\boldsymbol{\sigma}$, given that the state prior to measurement is $|0\rangle$. What is the state of the system after measurement if $+1$ is obtained?

Solution

Concepts Involved: Quantum Measurement, Projective Measurement.

The probability of obtaining the result $+1$ is given by:

$$p(+) = \langle 0|P_+|0\rangle = \langle 0|\frac{I + \mathbf{v}\cdot\boldsymbol{\sigma}}{2}|0\rangle$$

We recall from from Exercise 2.35 that:

$$\mathbf{v}\cdot\boldsymbol{\sigma} = \begin{bmatrix} v_3 & v_1 - iv_2 \\ v_1 + iv_2 & -v_3 \end{bmatrix} = v_3 |0\rangle\langle 0| + (v_1 - iv_2) |0\rangle\langle 1| + (v_1 + iv_2) |1\rangle\langle 0| - v_3 |1\rangle\langle 1| \,.$$

Hence computing $p(+)$, we get:

$$\begin{aligned} p(+) &= \langle 0| \left(\frac{1}{2}|0\rangle + \frac{1}{2}\left(v_3|0\rangle + (v_1 + iv_2)|1\rangle\right)\right) \\ &= \langle 0| \left(\frac{1+v_3}{2}|0\rangle + \frac{v_1 + iv_2}{2}|1\rangle\right) \\ &= \frac{1+v_3}{2}\langle 0|0\rangle + \frac{v_1+iv_2}{2}\langle 0|1\rangle = \frac{1+v_3}{2} \end{aligned}$$

so the probability of measuring the $+1$ outcome is $\frac{1+v_3}{2}$. The state after the measurement of the $+1$ outcome is given by:

$$|0\rangle \mapsto \frac{P_+|0\rangle}{\sqrt{p(+)}} = \frac{\frac{1+v_3}{2}|0\rangle + \frac{v_1+iv_2}{2}|1\rangle}{\sqrt{\frac{1+v_3}{2}}} = \frac{1}{\sqrt{2(1+v_3)}}\left((1+v_3)|0\rangle + (v_1+iv_2)|1\rangle\right)$$

$\square$

Exercise 2.62

Show that any measurement where the measurement operators and the POVM elements coincide is a projective measurement.

Solution

Concepts Involved: Quantum Measurement, Projective Measurement, POVM Measurement.

Suppose we have that the measurement operators M_m are equal to the POVM elements E_m. In this case, we have that:

$$M_m = E_m \equiv M_m^\dagger M_m$$

$M_m^\dagger M_m$ is positive by Exercise 2.25, so it follows that M_m is positive and hence Hermitian by Exercise 2.24. Hence, $M_m^\dagger = M_m$, and therefore:

$$M_m = M_m^\dagger M_m = M_m^2$$

From which we conclude that M_m are projective measurement operators. $\square$

Exercise 2.63

Suppose a measurement is described by measurement operators M_m. Show that there exist unitary operators U_m such that $M_m = U_m\sqrt{E_m}$, where E_m is the POVM associated to the measurement.

Solution

Concepts Involved: Quantum Measurement, POVM Measurement, Polar Decomposition.

Since M_m is a linear operator, by the left polar decomposition there exists unitary U such that:

$$M_m = U\sqrt{M_m^\dagger M_m} = U\sqrt{E_m},$$

where in the last equality we use that $M_m^\dagger M_m = E_m$. $\square$

Exercise 2.64

$(*)$ Suppose Bob is given a quantum state chosen from a set $|\psi_1\rangle, \ldots, |\psi_m\rangle$ of linearly independent states. Construct a POVM $\{E_1, E_2, \ldots, E_{m+1}\}$ such that if outcome E_i occurs, $1 \leq i \leq m$, then Bob knows with certainty that he was given the state $|\psi_i\rangle$. (The POVM must be such that $\langle\psi_i|E_i|\psi_i\rangle > 0$ for each i.)

Solution

Concepts Involved: POVM Measurement, Orthogonality

Let $\mathcal{H}$ be the Hilbert space where the given states lie, and let V be the m-dimensional subspace spanned by $|\psi_1\rangle, \ldots, |\psi_m\rangle$. For each $i \in \{1, \ldots, m\}$, let W_i be the subspace of V spanned by $\big\{|\psi_j\rangle : j \neq i\big\}$. Let $W_i^\perp$ be the orthogonal complement of W_i which consists of all states in $\mathcal{H}$ orthogonal to all states in W_i. We then have that any vector in V can be written as the sum of a vector in W_i and $W_i^\perp \cap V$ (see for example Theorem 6.47 in Axler's *Linear Algebra Done Right*). Therefore, for any $|\psi_i\rangle$ we can write:

$$|\psi_i\rangle = |w_i\rangle + |p_i\rangle$$

Where $|w_i\rangle \in W_i$ and $|p_i\rangle \in W_i^\perp \cap V$. Define $E_i = \frac{|p_i\rangle\langle p_i|}{m}$. By construction, we have that for any $|\psi\rangle \in \mathcal{H}$:

$$\langle\psi|E_i|\psi\rangle = \frac{\big|\langle\psi|p_i\rangle\big|^2}{m} \geq 0$$

so the E_is are positive are required. Furthermore, defining $E_{i+1} = I - \sum_{i=1}^m E_i$ we again see that for any $|\psi\rangle \in \mathcal{H}$:

$$\langle\psi|E_{i+1}|\psi\rangle = \langle\psi|I|\psi\rangle - \sum_{i=1}^m \langle\psi|E_i|\psi\rangle = 1 - \sum_{i=1}^m \langle\psi|E_i|\psi\rangle \geq 1 - \sum_{i=1}^m \frac{1}{m} = 0$$

so E_{i+1} is also positive as required. Finally, to see that the $E_1, \dots E_m$ have the desired properties, observe by construction that since $|p_i\rangle \in W_i^{\perp} \cap V$, it follows that $\langle \psi_j | p_i \rangle = 0$ for any $j \neq i$ (as the $|p_i\rangle$ will be orthogonal to all the vectors in $\{|\psi_j\rangle : j \neq i\}$ by construction). Calculating $\langle \psi_i | E_i | \psi_i \rangle$, we observe that:

$$\langle \psi_i | E_i | \psi_i \rangle = \left(\langle w_i | + \langle p_i | \right) \frac{|p_i\rangle\langle p_i|}{m} \left(|w_i\rangle + |p_i\rangle \right) = \frac{\left| \langle p_i | p_i \rangle \right|^2}{m} = \frac{1}{m} \geq 0$$

so if Bob measures E_i, he can be certain that he was given the state $|\psi_i\rangle$. ☐

Exercise 2.65

Express the states $(|0\rangle + |1\rangle)/\sqrt{2}$ and $(|0\rangle - |1\rangle)/\sqrt{2}$ is a basis in which they are *not* the same up to a relative phase shift.

Solution

Concepts Involved: Linear Algebra, Phase

Let us define our basis to be $|+\rangle := (|0\rangle + |1\rangle)/\sqrt{2}$ and $|-\rangle := (|0\rangle - |1\rangle)/\sqrt{2}$. Our two states are then just the basis vectors of this basis $(|+\rangle, |-\rangle)$ and are not the same up to relative phase shift. ☐

Exercise 2.66

Show that the average value of the observable $X_1 Z_2$ for a two qubit system measured in the state $(|00\rangle + |11\rangle)/\sqrt{2}$ is zero.

Solution

Concepts Involved: Quantum Measurement, Expectation, Composite Systems

Computing the expectation value of $X_1 Z_2$, we get:

$$\begin{aligned}
\langle X_1 Z_2 \rangle &= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}} \right) X_1 Z_2 \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}} \right) \\
&= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}} \right) \left(\frac{X_1 Z_2 |00\rangle + X_1 Z_2 |11\rangle}{\sqrt{2}} \right) \\
&= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}} \right) \left(\frac{|10\rangle - |01\rangle}{\sqrt{2}} \right) \\
&= \frac{1}{2} \left(\langle 00|10\rangle - \langle 00|01\rangle + \langle 11|10\rangle - \langle 11|01\rangle \right) \\
&= \frac{1}{2} (0 + 0 + 0 + 0) \\
&= 0.
\end{aligned}$$

☐

Exercise 2.67

Suppose V is a Hilbert space with a subspace W. Suppose $U : W \mapsto V$ is a linear operator which preserves inner products, that is, for any $|w_1\rangle$ and $|w_2\rangle$ in W,

$$\langle w_1|U^\dagger U|w_2\rangle = \langle w_1|w_2\rangle$$

Prove that there exists a unitary operator $U' : V \mapsto V$ which *extends* U. That is, $U'\,|w\rangle = U\,|w\rangle$ for all $|w\rangle$ in W, but U' is defined on the entire space V. Usually we omit the prime symbol $'$ and just write U to denote the extension.

Solution

Concepts Involved: Linear Algebra, Inner Products, Unitary Operators

---

By assumption we have that U is unitary on W as $\langle w_1|U^\dagger U|w_2\rangle = \langle w_1|w_2\rangle$ and hence $U^\dagger U = I_W$. Hence, it has spectral decomposition:

$$U = \sum_j \lambda_j \,|j\rangle\!\langle j|$$

where $\{|j\rangle\}$ is an orthonormal basis of the subspace W. Then, let $\{|j\rangle\} \cup \{|i\rangle\}$ be an orthnormal basis of the full space V. We then define:

$$U' = \sum_j \lambda_j \,|j\rangle\!\langle j| + \sum_i \,|i\rangle\!\langle i| = U + \sum_i \,|i\rangle\!\langle i|$$

We can then see that for any $|w\rangle \in W$ that:

$$U'|w\rangle = \left(U + \sum_i \,|i\rangle\!\langle i|\right)|w\rangle = U|w\rangle + \sum_j |i\rangle\langle i|w\rangle = U|w\rangle$$

where in the last line we use that $\langle i|w\rangle = 0$ as $|i\rangle$ are not in the subspace W. Finally, verifying the unitarity of U' we have that:

$$\begin{aligned}
U'^\dagger U' &= \left(\sum_j \lambda_j^* \,|j\rangle\!\langle j| + \sum_i \,|i\rangle\!\langle i|\right)\left(\sum_{j'} \lambda_j \,|j\rangle\!\langle j| + \sum_{i'} \,|i\rangle\!\langle i|\right) \\
&= \sum_j \sum_{j'} |j\rangle\langle j|j'\rangle\langle j'| + \sum_j \sum_{i'} |j\rangle\langle j|i'\rangle\langle i'| + \sum_i \sum_{j'} |i\rangle\langle i|j'\rangle\langle j'| + \sum_i \sum_{i'} |i\rangle\langle i|i'\rangle\langle i'| \\
&= \sum_j \sum_{j'} \langle j|j'\rangle \delta_{jj'} + \sum_i \sum_{i'} \langle i|i'\rangle \delta_{ii'} \\
&= \sum_j \,|j\rangle\!\langle j| + \sum_i \,|i\rangle\!\langle i| \\
&= I
\end{aligned}$$

$\square$

Exercise 2.68

Prove that $|\psi\rangle \neq |a\rangle|b\rangle$ for all single qubit states $|a\rangle$ and $|b\rangle$.

Solution

Concepts Involved: Linear Algebra, Composite Systems, Entanglement.
Recall that:

$$|\psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

Suppose for the take of contradiction that $|\psi\rangle = |a\rangle|b\rangle$ for some single qubit states $|a\rangle$ and $|b\rangle$. Then, we have that $|a\rangle = \alpha|0\rangle + \beta|1\rangle$ and $|b\rangle = \gamma|0\rangle + \delta|1\rangle$ for some $\alpha, \beta, \gamma, \delta \in \mathbb{C}$ such that $|\alpha|^2 + |\beta|^2 = 1$ and $|\gamma|^2 + |\delta|^2 = 1$. We then have that:

$$|a\rangle|b\rangle = (\alpha|0\rangle + \beta|1\rangle)(\gamma|0\rangle + \delta|1\rangle) = \alpha\gamma|00\rangle + \alpha\delta|01\rangle + \beta\gamma|10\rangle + \beta\delta|11\rangle.$$

Where we have used the linearity of the tensor product (though we supress the $\otimes$ symbols in the above expression). We then have that:

$$|\psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} = \alpha\gamma|00\rangle + \alpha\delta|01\rangle + \beta\gamma|10\rangle + \beta\delta|11\rangle$$

which forces $\alpha\delta = 0$ and $\beta\gamma = 0$. However, we then have that at least one of $\alpha\gamma$ or $\beta\delta$ is also zero, and we thus reach a contradiction. $\square$

Exercise 2.69

Verify that the Bell basis forms an orthonormal basis for the two qubit state space.

Solution

Concepts Involved: Linear Algebra, Orthogonality, Bases, Composite Systems.
Recall that the bell basis is given by:

$$|B_{00}\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \quad |B_{01}\rangle\frac{|00\rangle - |11\rangle}{\sqrt{2}}, \quad |B_{10}\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}, \quad |B_{11}\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$$

We first verify orthonormality. We observe that:

$$
\begin{aligned}
\langle B_{00}|B_{00}\rangle &= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right)\left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 00|00\rangle + \langle 00|11\rangle + \langle 11|00\rangle + \langle 11|11\rangle\right) = 1 \\
\langle B_{00}|B_{01}\rangle &= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right)\left(\frac{|00\rangle - |11\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 00|00\rangle - \langle 00|11\rangle + \langle 11|00\rangle - \langle 11|11\rangle\right) = 0 \\
\langle B_{00}|B_{10}\rangle &= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right)\left(\frac{|01\rangle + |10\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 00|01\rangle + \langle 00|10\rangle + \langle 11|01\rangle + \langle 11|10\rangle\right) = 0 \\
\langle B_{00}|B_{11}\rangle &= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right)\left(\frac{|01\rangle - |10\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 00|01\rangle - \langle 00|10\rangle + \langle 11|01\rangle - \langle 11|10\rangle\right) = 0 \\
\langle B_{01}|B_{01}\rangle &= \left(\frac{\langle 00| - \langle 11|}{\sqrt{2}}\right)\left(\frac{|00\rangle - |11\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 00|00\rangle - \langle 00|11\rangle - \langle 11|00\rangle + \langle 11|11\rangle\right) = 1 \\
\langle B_{01}|B_{10}\rangle &= \left(\frac{\langle 00| - \langle 11|}{\sqrt{2}}\right)\left(\frac{|01\rangle + |10\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 00|01\rangle + \langle 00|10\rangle - \langle 11|01\rangle - \langle 11|10\rangle\right) = 0 \\
\langle B_{01}|B_{11}\rangle &= \left(\frac{\langle 00| - \langle 11|}{\sqrt{2}}\right)\left(\frac{|01\rangle - |10\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 00|01\rangle - \langle 00|10\rangle - \langle 11|01\rangle + \langle 11|10\rangle\right) = 0 \\
\langle B_{10}|B_{10}\rangle &= \left(\frac{\langle 01| + \langle 10|}{\sqrt{2}}\right)\left(\frac{|01\rangle + |10\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 01|01\rangle + \langle 01|10\rangle + \langle 10|01\rangle + \langle 10|10\rangle\right) = 1 \\
\langle B_{10}|B_{11}\rangle &= \left(\frac{\langle 01| + \langle 10|}{\sqrt{2}}\right)\left(\frac{|01\rangle - |10\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 01|01\rangle - \langle 01|10\rangle + \langle 10|01\rangle - \langle 10|10\rangle\right) = 0 \\
\langle B_{11}|B_{11}\rangle &= \left(\frac{\langle 01| - \langle 10|}{\sqrt{2}}\right)\left(\frac{|01\rangle - |10\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 01|01\rangle - \langle 01|10\rangle - \langle 10|01\rangle + \langle 10|10\rangle\right) = 1
\end{aligned}
$$

so orthonormality is verified. We know show that it is a basis. Recall that we can write any vector $|\psi\rangle$ in the 2 qubit state space as:

$$|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$$

where $\alpha, \beta, \gamma, \delta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1$. We then observe that this is equivalent to:

$$\left(\frac{\alpha+\delta}{\sqrt{2}}\right)|B_{00}\rangle + \left(\frac{\alpha-\delta}{\sqrt{2}}\right)|B_{01}\rangle + \left(\frac{\beta+\gamma}{\sqrt{2}}\right)|B_{10}\rangle + \left(\frac{\beta-\gamma}{\sqrt{2}}\right)|B_{11}\rangle \quad (*)$$

as:

$$
\begin{aligned}
\left(\frac{\alpha+\delta}{\sqrt{2}}\right)\frac{|00\rangle+|11\rangle}{\sqrt{2}} + \left(\frac{\alpha-\delta}{\sqrt{2}}\right)\frac{|00\rangle-|11\rangle}{\sqrt{2}} &+ \left(\frac{\beta+\gamma}{\sqrt{2}}\right)\frac{|01\rangle+|10\rangle}{\sqrt{2}} + \left(\frac{\beta-\gamma}{\sqrt{2}}\right)\frac{|01\rangle-|10\rangle}{\sqrt{2}} \\
&= \left(\frac{\alpha}{2} + \frac{\alpha}{2} + \frac{\delta}{2} - \frac{\delta}{2}\right)|00\rangle + \left(\frac{\alpha}{2} - \frac{\alpha}{2} + \frac{\delta}{2} + \frac{\delta}{2}\right)|11\rangle \\
&+ \left(\frac{\beta}{2} + \frac{\beta}{2} + \frac{\gamma}{2} - \frac{\gamma}{2}\right)|01\rangle + \left(\frac{\beta}{2} - \frac{\beta}{2} + \frac{\gamma}{2} + \frac{\gamma}{2}\right)|10\rangle \\
&= \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle = |\psi\rangle
\end{aligned}
$$

Hence $(*)$ shows that the Bell states form a basis. $\square$

Exercise 2.70

Suppose E is any positive operator acting on Alice's qubit Show that $\langle\psi|E\otimes I|\psi\rangle$ *takes the same value* when $|\psi\rangle$ is any of the four Bell states. Suppose some malevolent third party ('Eve') intercepts Alice's qubit on the way to Bob in the superdense coding protocol. Can Eve infer anything about which of the four possible bit strings $00, 01, 10, 11$ Alice is trying to send? If so, how, or if not, why not?

Solution

Concepts Involved: Linear Algebra, Superdense Coding, Quantum Measurement

---

Let E be a positive operator. We have that E for a single qubit can be written as a linear combination of the Pauli matrices:

$$E = a_1 I + a_2 X + a_3 Y + a_4 Z$$

To see that this is the case, consider that the vector space of linear operators acting on a single qubit has dimension 4 (one easy way to see this is that the matrix representations of these operators have 4 entries). Hence, any set of 4 linearly independent linear operators form a basis for the space. As I, X, Y, Z are linearly independent, it follows that they form a basis of the space of linear operators on one qubit. Hence any E can be written as above (Remark: the above decomposition into Paulis is possible regardless of whether E is positive or not).
We then have that:

$$\begin{aligned}
\langle B_{00}|E\otimes I|B_{00}\rangle &= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right) E\otimes I\left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right)\\
&= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right)(a_1 I + a_2 X + a_3 Y + a_4 Z)\otimes I\left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right)\\
&= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right)\left(a_1\frac{|00\rangle + |11\rangle}{\sqrt{2}} + a_2\frac{|10\rangle + |01\rangle}{\sqrt{2}} + a_3\frac{i|10\rangle - i|01\rangle}{\sqrt{2}} + a_4\frac{|00\rangle - |11\rangle}{\sqrt{2}}\right)\\
&= \frac{1}{2}(a_1 + a_1 + a_4 - a_4)\\
&= a_1
\end{aligned}$$

where in the second last equality we use the orthonormality of $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. Repeating the same process for the other Bell states, we have:

$$\begin{aligned}
\langle B_{01}|E\otimes I|B_{01}\rangle &= \left(\frac{\langle 00| - \langle 11|}{\sqrt{2}}\right)(a_1 I + a_2 X + a_3 Y + a_4 Z)\otimes I\left(\frac{|00\rangle - |11\rangle}{\sqrt{2}}\right)\\
&= \left(\frac{\langle 00| - \langle 11|}{\sqrt{2}}\right)\left(a_1\frac{|00\rangle - |11\rangle}{\sqrt{2}} + a_2\frac{|10\rangle - |01\rangle}{\sqrt{2}} + a_3\frac{i|10\rangle + i|01\rangle}{\sqrt{2}} + a_4\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right)\\
&= \frac{1}{2}(a_1 + a_1 + a_4 - a_4)\\
&= a_1
\end{aligned}$$

$$
\begin{aligned}
\langle B_{10}|E \otimes I|B_{10}\rangle &= \left(\frac{\langle 01| + \langle 10|}{\sqrt{2}}\right)(a_1 I + a_2 X + a_3 Y + a_4 Z) \otimes I \left(\frac{|01\rangle + |10\rangle}{\sqrt{2}}\right) \\
&= \left(\frac{\langle 01| + \langle 10|}{\sqrt{2}}\right)\left(a_1\frac{|01\rangle + |10\rangle}{\sqrt{2}} + a_2\frac{|11\rangle + |00\rangle}{\sqrt{2}} + a_3\frac{i|11\rangle - i|00\rangle}{\sqrt{2}} + a_4\frac{|01\rangle - |10\rangle}{\sqrt{2}}\right) \\
&= \frac{1}{2}(a_1 + a_1 + a_4 - a_4) \\
&= a_1
\end{aligned}
$$

$$
\begin{aligned}
\langle B_{01}|E \otimes I|B_{01}\rangle &= \left(\frac{\langle 01| - \langle 10|}{\sqrt{2}}\right)(a_1 I + a_2 X + a_3 Y + a_4 Z) \otimes I \left(\frac{|01\rangle - |10\rangle}{\sqrt{2}}\right) \\
&= \left(\frac{\langle 01| - \langle 10|}{\sqrt{2}}\right)\left(a_1\frac{|01\rangle - |10\rangle}{\sqrt{2}} + a_2\frac{|11\rangle - |00\rangle}{\sqrt{2}} + a_3\frac{i|11\rangle + i|00\rangle}{\sqrt{2}} + a_4\frac{|01\rangle + |10\rangle}{\sqrt{2}}\right) \\
&= \frac{1}{2}(a_1 + a_1 + a_4 - a_4) \\
&= a_1
\end{aligned}
$$

Now, suppose that Eve intercepts Alice's qubit. Eve cannot infer anything about which of the four possible bit strings that Alice is trying to send, as any single-qubit measurement that Eve can perform on the intercepted qubit will return the value:

$$\langle\psi|M^\dagger M \otimes I|\psi\rangle$$

Where M is the (single-qubit) measurement operator. But, $M^\dagger M$ is positive, so by the above argument, the measurement outcome will be the same regardless of which Bell state $|\psi\rangle$ is. Hence, Eve cannot obtain the information about the bit string. $\square$

Exercise 2.71: Criterion to decide if a state is mixed or pure

Let ρ be a density operator. Show that $\mathrm{tr}\left(\rho^2\right) \leq 1$, with equality if and only if ρ is a pure state.

Solution

Concepts Involved: Linear Algebra, Trace, Density Operators, Pure States, Mixed States.
Recall that a density operator ρ is pure if:

$$\rho = |\psi\rangle\langle\psi|$$

for some normalized quantum state vector $|\psi\rangle$.

Since ρ is a positive operator, by the spectral decomposition we have that:

$$\rho = \sum_i p_i |i\rangle\langle i|$$

where $p_i \geq 0$ (due to positivity) and $|i\rangle$ are orthonormal. Furthermore, by the property of density operators, we have that $\mathrm{tr}(\rho) = 1$, hence:

$$\mathrm{tr}(\rho) = \mathrm{tr}\left(\sum_i p_i |i\rangle\langle i|\right) = \sum_i p_i \,\mathrm{tr}\big(|i\rangle\langle i|\big) = \sum_i p_i = 1$$

where in the second equality we use the linearity of the trace. We obtain that $0 \leq p_i \leq 1$ for each i. Calculating ρ^2, we have that:

$$\rho^2 = \left(\sum_i p_i |i\rangle\langle i|\right)\left(\sum_{i'} p_{i'} |i'\rangle\langle i'|\right) = \sum_i \sum_{i'} p_i p_{i'} |i\rangle\langle i|i'\rangle\langle i| = \sum_i \sum_{i'} p_i p_{i'} |i\rangle\langle i'| \,\delta_{ii'} = \sum_i p_i^2 |i\rangle\langle i|$$

Hence:

$$\mathrm{tr}\left(\rho^2\right) = \sum_i p_i^2 \,\mathrm{tr}\big(|i\rangle\langle i|\big) = \sum_i p_i^2 \leq \sum_i p_i = 1$$

where in the inequality we use the fact that $p_i^2 \leq p_i$ as $0 \leq p_i \leq 1$. The inequality becomes an equality when $p_i^2 = p_i$, that is, when $p_i = 0$ or $p_i = 1$. In order for $\mathrm{tr}(\rho) = 1$ to hold, we have that $p_i = 1$ for one i and zero for all others. Hence, ρ in this case is a pure state. Conversely, suppose ρ is a pure state. Then:

$$\mathrm{tr}\left(\rho^2\right) = \mathrm{tr}\big(|\psi\rangle\langle\psi|\psi\rangle\langle\psi|\big) = \mathrm{tr}\big(|\psi\rangle\langle\psi|\big) = 1.$$

□

Exercise 2.72: Bloch sphere for mixed states

The Bloch sphere picture for pure states of a single qubit was introduced in Section 1.2. This description has an important generalization to mixed states as follows.

(1) Show that an arbitrary density matrix for a mixed state qubit may be written as

$$\rho = \frac{I + \mathbf{r}\cdot\boldsymbol{\sigma}}{2},$$

Where $\mathbf{r}$ is a real three-dimensional vector such that $\|\mathbf{r}\| \leq 1$. This vector is known as the *Bloch vector* for the state ρ.

(2) What is the Bloch vector representation for the state $\rho = I/2$?

(3) Show that a state ρ is pure if and only if $\|\mathbf{r}\| = 1$.

(4) Show that for pure states the description of the Bloch vector we have given coincides with that in Section 1.2.

Solution

Concepts Involved: Linear Algebra, Trace, Density Operators, Pure States, Mixed States

(1) Since $\{I, X, Y, Z\}$ form an basis the vector space of single-qubit linear operators, we can write (for any ρ, regardless of whether it is a density operator or not):

$$\rho = a_1 I + a_2 X + a_3 Y + a_4 Z$$

for constants $a_1, a_2, a_3, a_4 \in \mathbb{C}$. Since ρ is a Hermitian operator, we find that each of these constants are actually real, as:

$$a_1 I + a_2 X + a_3 Y + a_4 Z = \rho = \rho^\dagger = a_1^* I^\dagger + a_2^* X^\dagger + a_3^* Y^\dagger + a_4^* Z^\dagger = a_1^* I + a_2^* X + a_3^* Y + a_4^* Z$$

Now, we require that $\operatorname{tr}(\rho) = 1$ for any density operator, hence:

$$\operatorname{tr}(\rho) = \operatorname{tr}(a_1 I + a_2 X + a_3 Y + a_4 Z) = a_1 \operatorname{tr}(I) + a_2 \operatorname{tr}(X) + a_3 \operatorname{tr}(Y) + a_4 \operatorname{tr}(Z) = 2a_1 = 1$$

from which we obtain that $a_1 = \frac{1}{2}$. Note that in the second equality we use the linearity of the trace, and in the third equality we use that $\operatorname{tr}(I) = 2$ and $\operatorname{tr}(\sigma_i) = 0$ for $i \in \{1, 2, 3\}$ (Exercise 2.36). Calculating ρ^2, we have that:

$$\begin{aligned}\rho^2 = \frac{1}{4} I &+ \frac{a_2}{2} X + \frac{a_3}{2} Y + \frac{a_4}{2} Z + \frac{a_2}{2} X + a_2^2 X^2 + a_2 a_3 XY + a_2 a_4 XZ \\ &+ \frac{a_3}{2} Y + a_3 a_2 YX + a_3^2 Y^2 + a_3 a_4 YZ + \frac{a_4}{2} Z + a_4 a_2 ZX + a_4 a_3 ZY + a_4^2 Z^2\end{aligned}$$

Now, using that $\left\{\sigma_i, \sigma_j\right\} = 0$ for $i, j \in \{1, 2, 3\}$, $i \neq j$ and that $\sigma_i^2 = I$ for any $i \in \{1, 2, 3\}$ (Exercise 2.41), the above simplifies to:

$$\rho^2 = \left(\frac{1}{4} + a_2^2 + a_3^2 + a_4^2\right) I + a_2 X + a_3 Y + a_4 Z$$

Taking the trace of ρ^2 we have that:

$$\operatorname{tr}\left(\rho^2\right) = \left(\frac{1}{4} + a_2^2 + a_3^2 + a_4^2\right) \operatorname{tr}(I) + a_2 \operatorname{tr}(X) + a_3 \operatorname{tr}(Y) + a_4 \operatorname{tr}(Z) = 2\left(\frac{1}{4} + a_2^2 + a_3^2 + a_4^2\right)$$

From the previous exercise (Exercise 2.71) we know that $\operatorname{tr}\left(\rho^2\right) \leq 1$, so:

$$2\left(\frac{1}{4} + a_2^2 + a_3^2 + a_4^2\right) \leq 1 \implies a_2^2 + a_3^2 + a_4^2 \leq \frac{1}{4} \implies \sqrt{a_2^2 + a_3^2 + a_4^2} \leq \frac{1}{2}$$

Hence we can write:

$$\rho = \frac{I + r_x X + r_y Y + r_z Z}{2} = \frac{I + \mathbf{r} \cdot \boldsymbol{\sigma}}{2}$$

with $\|\mathbf{r}\| \leq 1$.

(2) The Bloch representation for the state $\rho = \frac{I}{2}$ is the above form with $\mathbf{r} = \mathbf{0}$. This vector corresponds to the center of the Bloch sphere, which is a maximally mixed state ($\operatorname{tr}\left(\rho^2\right)$ is minimized, with $\operatorname{tr}\left(\rho^2\right) = \frac{1}{2}$).

(3) From the calculation in part (1), we know that for any ρ:

$$\operatorname{tr}\left(\rho^2\right) = 2\left(\frac{1 + r_x^2 + r_y^2 + r_z^2}{4}\right) = \frac{1 + r_x^2 + r_y^2 + r_z^2}{2}$$

if $\|\mathbf{r}\| = 1$, then $r_x^2 + r_y^2 + r_z^2 = 1$. Hence, $\operatorname{tr}\left(\rho^2\right) = 1$ and ρ is pure by Exercise 2.71. Conversely, suppose ρ is pure. Then, $\operatorname{tr}\left(\rho^2\right) = 1$, so:

$$\frac{1 + r_x^2 + r_y^2 + r_z^2}{2} = 1 \implies r_x^2 + r_y^2 + r_z^2 = 1 \implies \|\mathbf{r}\| = 1.$$

(4) In section 1.2, we looked at states that lie on the surface of the Bloch sphere, which we parameterized as:

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi}\sin\left(\frac{\theta}{2}\right)|1\rangle.$$

Calculating the density operator corresponding to $|\psi\rangle$, we have:

$$\begin{aligned}
\rho = |\psi\rangle\langle\psi| &= \cos^2\left(\frac{\theta}{2}\right)|0\rangle\langle 0| + \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)e^{-i\varphi}|0\rangle\langle 1| \\
&\quad + \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)e^{i\varphi}|1\rangle\langle 0| + \sin^2\left(\frac{\theta}{2}\right)|1\rangle\langle 1| \\
&= \cos^2\left(\frac{\theta}{2}\right)|0\rangle\langle 0| + \frac{\sin(\theta)e^{-i\varphi}}{2}|0\rangle\langle 1| + \frac{\sin(\theta)e^{i\varphi}}{2}|1\rangle\langle 0| + \sin^2\left(\frac{\theta}{2}\right)|1\rangle\langle 1|
\end{aligned}$$

Conversely, we have that (in the computational basis) our proposed form of $\rho = \frac{I + \mathbf{r}\cdot\boldsymbol{\sigma}}{2}$ can be represented as:

$$\rho = \frac{1 + r_z}{2}|0\rangle\langle 0| + \frac{r_x - ir_y}{2}|0\rangle\langle 1| + \frac{r_x + ir_y}{2}|1\rangle\langle 0| + \frac{1 - r_z}{2}|1\rangle\langle 1|$$

Solving for r_x, r_y, r_z by equating the two expressions for ρ (using Euler's formula and $\sin(2\theta) = 2\cos(\theta)\sin(\theta)$), we have:

$$r_x = \cos(\varphi)\sin(\theta), \quad r_y = \sin(\varphi)\sin(\theta), \quad r_z = 2\cos^2\left(\frac{\theta}{2}\right) - 1 = \cos(\theta)$$

Calculating $\|\mathbf{r}\|$ we have that:

$$\begin{aligned}
\|\mathbf{r}\| = \sqrt{r_x^2 + r_y^2 + r_z^2} &= \sqrt{\cos^2(\varphi)\sin^2(\theta) + \sin^2(\varphi)\cos^2(\theta) + \cos^2(\theta)} \\
&= \sqrt{\sin^2(\theta) + \cos^2(\theta)} \\
&= 1
\end{aligned}$$

so we see that indeed, the two definitions coincide for pure states (as $\|\mathbf{r}\| = 1$).

$\square$

Exercise 2.73

$(*)$ Let ρ be a density operator. A *minimal ensemble* for ρ is an ensemble $\{p_i,\, |\psi_i\rangle\}$ containing a number of elements equal to the rank of ρ. Let $|\psi\rangle$ be any state in the support of ρ. (The *support* of a Hermitian operator A is the vector space spanned by the eigenvectors of A with non-zero eigenvalues.) Show that there is a minimal ensemble for ρ that contains $|\psi\rangle$, and moreover that in any such ensemble $|\psi\rangle$ must appear with probability

$$p_i = \frac{1}{\langle\psi_i|\,\rho^{-1}\,|\psi_i\rangle},$$

where ρ^{-1} is defined to be the inverse of ρ, when ρ is considered as an operator acting only on the support of ρ. (This definition removes the problem that ρ may not have an inverse.)

Solution

Concepts Involved: Below, we will use the unitary freedom in the ensemble for density matrices which is also known as Uhlmann's theorem. Specifically recall that $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i| = \sum_j q_j|\varphi_j\rangle\langle\varphi_j|$ for ensembles $\{p_i, |\psi_i\rangle\}$ and $\{q_j, |\varphi_j\rangle\}$ if and only if

$$\sqrt{p_i}\,|\psi_i\rangle = \sum_j u_{ij}\sqrt{q_j}\,|\varphi_j\rangle$$

for some unitary matrix u_{ij}.

Using the spectral decomposition of the density matrix we have

$$\rho = \sum_{k=1}^{r}\lambda_k|k\rangle\langle k| \quad \text{with } \lambda_k > 0$$

where all the eigenvectors with eigenvalue 0 have been removed. Thus, the set of vectors $S = \{|k\rangle\}_{k=1}^{r}$ forms a spanning set set for the support of ρ. An element in the support of ρ can thus be decomposed as

$$|\psi_i\rangle = \sum_k c_{ik}|k\rangle = \sum_k \langle k|\psi_i\rangle|k\rangle$$

Assuming that $|\psi_i\rangle$ occurs with probability p_i, we can use the Uhlmann's theorem quoted above to arrive at the relation

$$\sqrt{p_i}|\psi_i\rangle \overset{?}{=} \sum_k u_{ik}\sqrt{\lambda_k}|k\rangle = \sqrt{p_i}\sum_k\langle k|\psi_i\rangle|k\rangle,$$

which allows us relate the elements of one of the columns (i th) of the unitary matrix to

$$u_{ik}\sqrt{\lambda_k} \overset{?}{=} \sqrt{p_i}\langle k|\psi_i\rangle.$$

Such a relation can always be satisfied for a unitary matrix with dimension r. As u is unitary, we have

$$\begin{aligned}\sum_k |u_{ik}|^2 = 1 \implies 1 &= \sum_k \frac{p_i\langle\psi_i|k\rangle\langle k|\psi_i\rangle}{\lambda_k} \\ \implies p_i &= \sum_k \frac{\lambda_k}{\langle\psi_i|k\rangle\langle k|\psi_i\rangle} \\ &= \frac{1}{\langle\psi_i|\sum_k \frac{1}{\lambda_k}|k\rangle\langle k|\psi_i\rangle} \\ &= \frac{1}{\langle\psi_i|\rho^{-1}|\psi_i\rangle}\ .\end{aligned}$$

$$\begin{aligned}|\psi_j\rangle &= \rho\rho^{-1}|\psi_i\rangle \\ :&= \sum_{i=1}^{r} p_i|\psi_i\rangle\langle\psi_i|\rho^{-1}|\psi_j\rangle \\ &= \sum_{i=1}^{r} p_i\langle\psi_i|\rho^{-1}|\psi_j\rangle|\psi_i\rangle.\end{aligned}$$

But now note that $\{|\psi_j\rangle\}_{j=1}^{r}$ are linearly independent and $|\psi_j\rangle := \sum_{i=1}^{r}\delta_{ij}|\psi_i\rangle$.

$$\implies p_i\langle\psi_i|\rho^{-1}|\psi_i\rangle = 1.$$

Thus, the probability associated with the state $|\psi_i\rangle$ in the ensemble is given by

$$p_i = \frac{1}{\langle\psi_i|\rho^{-1}|\psi_i\rangle}\ .$$

□

Remark:

Exercise 2.74

Suppose a composite of systems A and B is in the state $|a\rangle\ |b\rangle$, where $|a\rangle$ is a pure state of system A, and $|b\rangle$ is a pure state of system B. Show that the reduced density operator of system A alone in a pure state.

Solution

Concepts Involved: Linear Algebra, Density Operators, Reduced Density Operators, Partial Trace, Pure States.

Suppose we have $|a\rangle|b\rangle \in A\otimes B$. Then, the density operator of the combined system is given as $\rho^{AB} = (|a\rangle|b\rangle)(\langle a|\langle b|) = |a\rangle\langle a|\otimes|b\rangle\langle b|$. Calculating the reduced density operator of system A by tracing out system B, we have

$$\rho^A = \mathrm{tr}_B(\rho_{AB}) = \mathrm{tr}_B(\,|a\rangle\langle a| \otimes |b\rangle\langle b|) = |a\rangle\langle a|\,\mathrm{tr}\big(\,|b\rangle\langle b|\big) = |a\rangle\langle a|\,\langle b|b\rangle = |a\rangle\langle a|\,.$$

Hence we find that $\rho^A = |a\rangle\langle a|$ is indeed a pure state. □

Exercise 2.75

For each of the four Bell states, find the reduced density operator for each qubit.

Solution

Concepts Involved: Linear Algebra, Density Operators, Reduced Density Operators, Partial Trace.

For the bell state $|B_{00}\rangle$, we have the density operator:

$$\rho = \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right)\left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right) = \frac{|00\rangle\langle 00| + |11\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 11|}{2}$$

Obtaining the reduced density operator for qubit A, we have:

$$\begin{aligned}\rho^A = \mathrm{tr}_B(\rho) &= \frac{\mathrm{tr}_B(\,|00\rangle\langle 00|) + \mathrm{tr}_B(\,|00\rangle\langle 11|) + \mathrm{tr}_B(\,|11\rangle\langle 00|) + \mathrm{tr}_B(\,|11\rangle\langle 11|)}{2}\\ &= \frac{|0\rangle\langle 0|\,\langle 0|0\rangle + |0\rangle\langle 1|\,\langle 1|0\rangle + |1\rangle\langle 0|\,\langle 0|1\rangle + |1\rangle\langle 1|\,\langle 1|1\rangle}{2}\\ &= \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2}\\ &= \frac{I}{2}\end{aligned}$$

Obtaining the reduced density operator for qubit B, we have:

$$\begin{aligned}\rho^B = \mathrm{tr}_A(\rho) &= \frac{\mathrm{tr}_A(\,|00\rangle\langle 00|) + \mathrm{tr}_A(\,|00\rangle\langle 11|) + \mathrm{tr}_A(\,|11\rangle\langle 00|) + \mathrm{tr}_A(\,|11\rangle\langle 11|)}{2}\\ &= \frac{\langle 0|0\rangle\,|0\rangle\langle 0| + \langle 1|0\rangle\,|0\rangle\langle 1| + \langle 0|1\rangle\,|1\rangle\langle 0| + \langle 1|1\rangle\,|1\rangle\langle 1|}{2}\\ &= \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2}\\ &= \frac{I}{2}\end{aligned}$$

We repeat a similar process for the other four bell states. For $|B_{01}\rangle$, we have:

$$\begin{aligned}
\rho &= \left(\frac{|00\rangle - |11\rangle}{\sqrt{2}}\right)\left(\frac{\langle 00| - \langle 11|}{\sqrt{2}}\right) = \frac{|00\rangle\langle 00| - |00\rangle\langle 11| - |11\rangle\langle 00| + |11\rangle\langle 11|}{2} \\
\rho^A &= \frac{\mathrm{tr}_B(|00\rangle\langle 00|) - \mathrm{tr}_B(|00\rangle\langle 11|) - \mathrm{tr}_B(|11\rangle\langle 00|) + \mathrm{tr}_B(|11\rangle\langle 11|)}{2} \\
&= \frac{|0\rangle\langle 0| \langle 0|0\rangle - |0\rangle\langle 1| \langle 1|0\rangle - |1\rangle\langle 0| \langle 0|1\rangle + |1\rangle\langle 1| \langle 1|1\rangle}{2} \\
&= \frac{I}{2} \\
\rho^B &= \frac{\mathrm{tr}_A(|00\rangle\langle 00|) - \mathrm{tr}_A(|00\rangle\langle 11|) - \mathrm{tr}_A(|11\rangle\langle 00|) + \mathrm{tr}_A(|11\rangle\langle 11|)}{2} \\
&= \frac{\langle 0|0\rangle |0\rangle\langle 0| + \langle 1|0\rangle |0\rangle\langle 1| + \langle 0|1\rangle |1\rangle\langle 0| + \langle 1|1\rangle |1\rangle\langle 1|}{2} \\
&= \frac{I}{2}
\end{aligned}$$

For $|B_{10}\rangle$, we have:

$$\begin{aligned}
\rho &= \left(\frac{|01\rangle + |10\rangle}{\sqrt{2}}\right)\left(\frac{\langle 01| + \langle 10|}{\sqrt{2}}\right) = \frac{|01\rangle\langle 01| + |01\rangle\langle 10| + |10\rangle\langle 01| + |10\rangle\langle 10|}{2} \\
\rho^A &= \frac{\mathrm{tr}_B(|01\rangle\langle 01|) + \mathrm{tr}_B(|01\rangle\langle 10|) + \mathrm{tr}_B(|10\rangle\langle 01|) + \mathrm{tr}_B(|10\rangle\langle 10|)}{2} \\
&= \frac{|0\rangle\langle 0| \langle 1|+\rangle |0\rangle\langle 1| \langle 0|1\rangle + |1\rangle\langle 0| \langle 1|0\rangle + |1\rangle\langle 1| \langle 0|0\rangle}{2} \\
&= \frac{I}{2} \\
\rho^B &= \frac{\mathrm{tr}_A(|01\rangle\langle 01|) + \mathrm{tr}_A(|01\rangle\langle 10|) + \mathrm{tr}_A(|10\rangle\langle 01|) + \mathrm{tr}_A(|10\rangle\langle 10|)}{2} \\
&= \frac{\langle 0|0\rangle |1\rangle\langle 1| + \langle 1|0\rangle |1\rangle\langle 0| + \langle 0|1\rangle |0\rangle\langle 1| + \langle 1|1\rangle |1\rangle\langle 1|}{2} \\
&= \frac{I}{2}
\end{aligned}$$

Finally, for $|B_{11}\rangle$ we have:

$$\rho = \left(\frac{|01\rangle - |10\rangle}{\sqrt{2}}\right)\left(\frac{\langle 01| - \langle 10|}{\sqrt{2}}\right) = \frac{|01\rangle\langle 01| - |01\rangle\langle 10| - |10\rangle\langle 01| + |10\rangle\langle 10|}{2}$$

$$\begin{aligned}
\rho^A &= \frac{\mathrm{tr}_B(|01\rangle\langle 01|) - \mathrm{tr}_B(|01\rangle\langle 10|) - \mathrm{tr}_B(|10\rangle\langle 01|) + \mathrm{tr}_B(|10\rangle\langle 10|)}{2} \\
&= \frac{|0\rangle\langle 0|\langle 1|-\rangle\,|0\rangle\langle 1|\langle 0|1\rangle - |1\rangle\langle 0|\langle 1|0\rangle + |1\rangle\langle 1|\langle 0|0\rangle}{2} \\
&= \frac{I}{2} \\
\rho^B &= \frac{\mathrm{tr}_A(|01\rangle\langle 01|) - \mathrm{tr}_A(|01\rangle\langle 10|) - \mathrm{tr}_A(|10\rangle\langle 01|) + \mathrm{tr}_A(|10\rangle\langle 10|)}{2} \\
&= \frac{\langle 0|0\rangle\,|1\rangle\langle 1| - \langle 1|0\rangle\,|1\rangle\langle 0| - \langle 0|1\rangle\,|0\rangle\langle 1| + \langle 1|1\rangle\,|1\rangle\langle 1|}{2} \\
&= \frac{I}{2}
\end{aligned}$$

□

Exercise 2.76

Extend the proof of the Schmidt decomposition to the case where A and B may have state space of different dimensionality.

Solution

Concepts Involved: Schmidt Decomposition, Singular Value Decomposition.
Note that for this problem we will use a more general form of the Singular Value Decomposition than proven in Nielsen and Chuang (that may have been encountered in a linear algebra course). Given an arbitrary $m \times n$ rectangular matrix A, there exists an $m \times m$ unitary matrix U and $n \times n$ unitary matrix V such that $A = U\Sigma V$ where Σ is a $m \times n$ rectangular diagonal matrix with non-negative reals on the diagonal (see `https://en.wikipedia.org/wiki/Singular_value_decomposition`).

Let $|m\rangle, |n\rangle$ be orthonormal bases for A and B. We can then write:

$$A = \sum_{mn} a_{mn}|m\rangle|n\rangle$$

for some $m \times n$ matrix of complex numbers a. Using the generalized SVD, we can write:

$$A = \sum_{min} u_{mi}d_{ii}v_{in}|m\rangle|n\rangle$$

where d_{ii} is a rectangular diagonal matrix. We can then define $|i_A\rangle = \sum_m u_{mi}\,|m\rangle$, $|i_B\rangle = \sum_n u_{in}|n\rangle$, and $\lambda_i = d_{ii}$ to yield the Schmidt decomposition. Note that we take $i = \min(m, n)$ and our sum only has as many terms as the dimensionality of the smaller space. □

Exercise 2.77

($*$) Suppose ABC is a three component quantum system. Show by example that there are quantum states $|\psi\rangle$ of such systems which can not be written in the form

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle|i_B\rangle|i_C\rangle$$

where λ_i are real numbers, and $|i_A\rangle, |i_B\rangle, |i_C\rangle$ are orthonormal bases of the respective systems.

Solution

Concepts Involved: Linear Algebra, Schmidt Decomposition.

Consider the state:

$$|\psi\rangle = |0\rangle \otimes |B_{00}\rangle = \frac{|000\rangle + |011\rangle}{\sqrt{2}}$$

we claim that this state cannot be written in the form:

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle|i_B\rangle|i_C\rangle$$

for orthonormal bases $|i_A\rangle, |i_B\rangle, |i_C\rangle$. Suppose for the sake of contradiction that we could write it in this form. We then make the observation that:

$$\begin{aligned}
\rho^A &= \mathrm{tr}_{BC}(\,|\psi\rangle\langle\psi|) = \sum_i \lambda_i^2\, |i_A\rangle\langle i_A| \\
\rho^B &= \mathrm{tr}_{AC}(\,|\psi\rangle\langle\psi|) = \sum_i \lambda_i^2\, |i_B\rangle\langle i_B| \\
\rho^C &= \mathrm{tr}_{AB}(\,|\psi\rangle\langle\psi|) = \sum_i \lambda_i^2\, |i_C\rangle\langle i_C|\,.
\end{aligned}$$

From this, we conclude that if it is possible to write $|\psi\rangle$ in such a form, then the eigenvalues of the reduced density matrices must all agree and be equal to λ_i^2. Computing the density matrix of the proposted $|\psi\rangle = |0\rangle \otimes |B_{00}\rangle$, we have:

$$\rho = \frac{|000\rangle\langle 000| + |000\rangle\langle 011| + |011\rangle\langle 000| + |011\rangle\langle 011|}{2}$$

Computing the reduced density matrices ρ^A and ρ^B, we find that:

$$\rho^A = \mathrm{tr}_{BC}(\rho) = |0\rangle\langle 0|$$

$$\rho^B = \mathrm{tr}_{AC}(\rho) = \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2}.$$

However, the former reduced density matrix has eigenvalues $\lambda_1^2 = 1, \lambda_2^2 = 0$, and the latter has $\lambda_1^2 = \frac{1}{2}$, $\lambda_2^2 = \frac{1}{2}$. This contradicts the fact that the λ_i^2s must match. $\square$

Remark: Necessary and Sufficient conditions for the tripartite (and higher order) Schmidt decompositions can be found here `https://arxiv.org/pdf/quant-ph/9504006.pdf`.

Exercise 2.78

Prove that a state $|\psi\rangle$ of a composite system AB is a product state if and only if it has a Schmidt number 1. Prove that $|\psi\rangle$ is a product state if and only if ρ^A (and thus ρ^B) are pure states.

Solution

Concepts Involved: Linear Algebra, Schmidt Decomposition, Schmidt Number, Reduced Density Operators.

Suppose $|\psi\rangle$ is a product state. Then, $|\psi\rangle = |0_A\rangle|0_B\rangle$ for some $|0_A\rangle, |0_B\rangle$, and we therefore have that $|\psi\rangle$ has Schmidt number 1 (it is already written in Schmidt decomposition form, and has one nonzero λ). Conversely, suppose $|\psi\rangle$ has Schmidt number 1. Then, $|\psi\rangle = 1|0_A\rangle|0_B\rangle + 0|1_A\rangle|1_B\rangle$ when writing $|\psi\rangle$ in its Schmidt decomposition. Therefore, $|\psi\rangle = |i_A\rangle|i_B\rangle$ and $|\psi\rangle$ is a product state.
Next, take any $|\psi\rangle$ and write out its Schmidt decomposition. We then get:

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle|i_B\rangle.$$

Hence:

$$\rho = \sum_i \lambda_i^2 \, |i_A\rangle\langle i_A| \otimes |i_B\rangle\langle i_B| \, .$$

Taking the partial trace of ρ to obtain ρ^A, we have:

$$\rho^A = \mathrm{tr}_B(\rho) = \sum_i \lambda_i^2 \,\mathrm{tr}_B\big(|i_A\rangle\langle i_A| \otimes |i_B\rangle\langle i_B|\big) = \sum_i \lambda_i^2 \, |i_A\rangle\langle i_A| \,\mathrm{tr}\big(|i_B\rangle\langle i_B|\big) = \sum_i \lambda_i^2 \, |i_A\rangle\langle i_A| \, .$$

Identically:

$$\rho^B = \mathrm{tr}_A(\rho) = \sum_i \lambda_i^2 \, |i_B\rangle\langle i_B| \, .$$

Now, suppose that $|\psi\rangle$ is a product state. Then, $|\psi\rangle$ has Schmidt number 1. Hence, only one of λ_1, λ_2 is nonzero. Hence, $\rho^A = |i_A\rangle\langle i_A|$ and $\rho^B = |i_B\rangle\langle i_B|$, so ρ^A, ρ^B are pure. Conversely, suppose ρ^A, ρ^B are pure. Then, we have that $\rho^A = |i_A\rangle\langle i_A|$ and $\rho^B = |i_B\rangle\langle i_B|$, so it follows that one of λ_1, λ_2 in the above equations for ρ^A, ρ^B must be zero. Therefore, $|\psi\rangle$ has Schmidt number 1, and is hence a product state. $\square$

Exercise 2.79

Consider a composite system consisting of two qubits. Find the Schmidt decomposition of the states

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}}; \quad \frac{|00\rangle + |01\rangle + |10\rangle + |11\rangle}{2}; \text{ and } \frac{|00\rangle + |01\rangle + |10\rangle}{\sqrt{3}}$$

Solution

Concepts Involved: Linear Algebra, Schmidt Decomposition, Reduced Density Matrices, Partial Trace.

For the first two expressions, by inspection we find that:

$$\begin{aligned}\frac{|00\rangle + |11\rangle}{\sqrt{2}} &= \frac{1}{\sqrt{2}}|0\rangle|0\rangle + \frac{1}{\sqrt{2}}|1\rangle|1\rangle \\ \frac{|00\rangle + |01\rangle + |10\rangle + |11\rangle}{2} &= 1|+\rangle|+\rangle + 0|-\rangle|-\rangle\end{aligned}$$

For the third expression, we require a little more work. We first note that the existence of the Schmidt decomposition guarantees that the state $|\psi\rangle = \frac{|00\rangle+|01\rangle+|10\rangle}{\sqrt{3}}$ can be written in the form $|\psi\rangle = \sum_{i=1}^{2}\lambda_i|i_A\rangle|i_B\rangle$ for some choice of orthonormal bases $|i_A\rangle, |i_B\rangle$. By the definition of reduced density matrices/partial trace, we can make the observation that:

$$\rho^A = \mathrm{tr}_B(\rho) = \mathrm{tr}_B(\,|\psi\rangle\langle\psi|) = \sum_{i=1}^{2}\lambda_i^2\,|i_A\rangle\langle i_A|\,\mathrm{tr}_B(\,|i_B\rangle\langle i_B|) = \sum_{i=1}^{2}\lambda_i^2\,|i_A\rangle\langle i_A|$$

and similarly that $\rho^B = \sum_{i=1}^{2}\lambda_i^2\,|i_B\rangle\langle i_B|$. Hence, to find the Schmidt decomposition of $|\psi\rangle$, we can compute the reduced density matrices and then solve for their eigenvalues λ_i^2 and eigenvectors $|i\rangle$. First solving for ρ, we have:

$$\rho = \frac{|00\rangle\langle 00| + |00\rangle\langle 01| + |00\rangle\langle 10| + |01\rangle\langle 00| + |01\rangle\langle 01| + |01\rangle\langle 10| + |10\rangle\langle 00| + |10\rangle\langle 01| + |10\rangle\langle 10|}{3}$$

Solving for the reduced density matrix ρ^A we have:

$$\rho^A = \mathrm{tr}_B(\rho) = \frac{2\,|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 1|}{3} \cong \frac{1}{3}\begin{bmatrix}2 & 1\\ 1 & 1\end{bmatrix}$$

Solving for the eigenvalues and (normalized) eigenvectors, we have:

$$\begin{aligned}\lambda_1^2 &= \frac{3+\sqrt{5}}{6}, |1_A\rangle = \frac{1}{\sqrt{10+2\sqrt{5}}}\left((1+\sqrt{5})|0\rangle + 2|1\rangle\right)\\ \lambda_2^2 &= \frac{3-\sqrt{5}}{6}, |2_A\rangle = \frac{1}{\sqrt{10-2\sqrt{5}}}\left((1-\sqrt{5})|0\rangle + 2|1\rangle\right).\end{aligned}$$

Next solving for the reduced density matrix ρ^B, we have:

$$\rho^B = \mathrm{tr}_A(\rho) = \frac{2\,|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 1|}{3} \cong \frac{1}{3}\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

This (of course) has the same eigenvectors and eigenvalues:

$$\lambda_1^2 = \frac{3+\sqrt{5}}{6}, |1_B\rangle = \frac{1}{\sqrt{10+2\sqrt{5}}}\left((1+\sqrt{5})|0\rangle + 2|1\rangle\right)$$
$$\lambda_2^2 = \frac{3-\sqrt{5}}{6}, |2_B\rangle = \frac{1}{\sqrt{10-2\sqrt{5}}}\left((1-\sqrt{5})|0\rangle + 2|1\rangle\right).$$

Hence the Schmidt decomposition of $|\psi\rangle$ is given by:

$$|\psi\rangle = \lambda_1|1_A\rangle|1_B\rangle + \lambda_2|2_A\rangle|2_B\rangle$$

where the expressions for the eigenvalues/eigenvectors are given above. □

Exercise 2.80

Suppose $|\psi\rangle$ and $|\varphi\rangle$ are two pure states of a composite quantum system with components A and B, with identical Schmidt coefficients. Show that there are unitary transformations U on a system A and V on system B such that $|\psi\rangle = (U \otimes V)|\varphi\rangle$.

Solution

Concepts Involved: Linear Algebra, Schmidt Decomposition, Unitary Operators.

We first prove a Lemma. Suppose we have two (orthonormal) bases $\{|i\rangle\}, \{|i'\rangle\}$ of a (n-dimensional) vector space A. We claim that the change of basis transformation U where $|i'\rangle = U|i\rangle$ is unitary.
To see this is the case, let $U = \sum_i |i'\rangle\langle i|$. By orthonormality, we see that $U|i\rangle = |i'\rangle$ as desired. Computing $U^\dagger$, we have that $U^\dagger = \sum_i(|i'\rangle\langle i|)^\dagger = \sum_i |i\rangle\langle i'|$. By orthonormality, we then see that $U^\dagger U = \sum_i |i\rangle\langle i| = I$ and hence U is unitary.
We now move onto the actual problem. By assumption, we can write $|\psi\rangle = \sum_i \lambda_i|i_A\rangle|i_B\rangle$ and $|\varphi\rangle = \sum_j \lambda_j|j_A\rangle|j_B\rangle$ where $\lambda_i = \lambda_j$ if $i = j$. By the lemma, there exists unitary change-of-basis matrices U, V such that $|i_A\rangle = U|j_A\rangle$ and $|i_B\rangle = V|j_B\rangle$. Hence, we have that:

$$|\psi\rangle = \sum_i \lambda_i|i_A\rangle|i_B\rangle = \sum_j \lambda_j(U|j_A\rangle)(V|j_B\rangle) = (U \otimes V)\sum_j \lambda_j|j_A\rangle|j_B\rangle = (U \otimes V)|\varphi\rangle$$

which is what we wanted to prove. □

Exercise 2.81: Freedom in purifications

Let $|AR_1\rangle$ and $|AR_2\rangle$ be two purifications of a state ρ^A to a composite system AR. Prove that there exists a unitary transformation U_R acting on system R such that $|AR_1\rangle = (I_A \otimes U_R)|AR_2\rangle$.

Solution

Concepts Involved: Linear Algebra, Schmidt Decomposition, Purification, Unitary Operators

Let $|AR_1\rangle, |AR_2\rangle$ be two purifications of ρ^A to a composite system AR. We can write the orthonormal decomposition of ρ^A as $\rho^A = \sum_i p_i \, |i^A\rangle\langle i^A|$, from which it follows that we can write:

$$|AR_1\rangle = \sum_i \sqrt{p_i}|i^A\rangle|i^R\rangle$$
$$|AR_2\rangle = \sum_i \sqrt{p_i}|i^A\rangle|i'^R\rangle$$

for some bases $\{|i\rangle\}, \{|i'\rangle\}$ of R. By the Lemma proven in the previous exercise, the transformation U_R such that $|i\rangle = U_R|i'\rangle$ is unitary, so hence:

$$\begin{aligned}|AR_1\rangle = \sum_i \sqrt{p_i}|i^A\rangle|i^R\rangle = \sum_i \sqrt{p_i}|i^A\rangle(U_R|i'^R\rangle) &= \sum_i \sqrt{p_i}(I_A|i^A\rangle)(U_R|i'^R\rangle) \\ &= (I_A \otimes U_R)\sum_i \sqrt{p_i}|i^A\rangle|i'^R\rangle \\ &= (I_A \otimes U_R)|AR_2\rangle\end{aligned}$$

which proves the claim. □

Exercise 2.82

Suppose $\{p_i, |\psi_i\rangle\}$ is an ensemble of states generating a density matrix $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|$ for a quantum system A. Introduce a system R with orthonormal basis $|i\rangle$.

(1) Show that $\sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle$ is a purification of ρ.

(2) Suppose we measure R in the basis $|i\rangle$, obtained outcome i. With what probability do we obtain the result i, and what is the corresponding state of system A?

(3) Let $|AR\rangle$ be *any* purification of ρ to the system AR. Show that there exists an orthonormal basis $|i\rangle$ in which R can be measured such that the corresponding post-measurement state for system A is $|\psi_i\rangle$ with probability p_i.

Solution

Concepts Involved: Linear Algebra, Purification, Schmidt Decomposition.

(1) To verify that $\sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle$ is a purification, we see that:

$$\begin{aligned}
\mathrm{tr}_R\left(\left(\sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle\right)\left(\sum_j \sqrt{p_j}\langle\psi_j|\langle j|\right)\right) &= \sum_i\sum_j \sqrt{p_ip_j}\,|\psi_i\rangle\!\langle\psi_j|\,\mathrm{tr}_R(|i\rangle\langle j|) \\
&= \sum_i\sum_j \sqrt{p_ip_j}|\psi_i\rangle\langle\psi_j|\delta_{ij} \\
&= \sum_i \sqrt{p_i^2}\,|\psi_i\rangle\!\langle\psi_i| \\
&= \sum_i p_i\,|\psi_i\rangle\!\langle\psi_i| \\
&= \rho
\end{aligned}$$

(2) We measure the observable $M_i = I_A\otimes\sum_i P_i = I_A\otimes\sum_i |i\rangle\!\langle i|$. The probability of obtaining outcome i is given by $p(i) = \langle AR|(I_A\otimes P_i)|AR\rangle$ (where $|AR\rangle = \sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle$), which we can calculate to be:

$$\begin{aligned}
p(i) &= \langle AR|(I_A\otimes |i\rangle\!\langle i|)|AR\rangle \\
&= \left(\sum_j \sqrt{p_j}\langle\psi_j|\langle j|\right)(I_A\otimes |i\rangle\!\langle i|)\left(\sum_k \sqrt{p_k}|\psi_k\rangle|k\rangle\right) \\
&= \sum_j\sum_j \sqrt{p_j}\sqrt{p_k}\langle\psi_j|\psi_k\rangle\delta_{ji}\delta_{ik} \\
&= p_i
\end{aligned}$$

The post measurement state is given by:

$$\begin{aligned}
\frac{(I_A\otimes P_i)|AR\rangle}{\sqrt{p(i)}} &= \frac{(I_A\otimes |i\rangle\!\langle i|)\sum_j \sqrt{p_j}|\psi_j\rangle|j\rangle}{\sqrt{p_i}} \\
&= \frac{\sum_j \sqrt{p_j}|\psi_j\rangle|j\rangle\delta_{ij}}{\sqrt{p_i}} \\
&= \frac{\sqrt{p_i}|\psi_i\rangle|i\rangle}{\sqrt{p_i}} \\
&= |\psi_i\rangle|i\rangle
\end{aligned}$$

so the corresponding state of system A is $|\psi_i\rangle$.

(3) Let $|AR\rangle$ be any purification of ρ to the combined system AR. We then have that $|AR\rangle$ has Schmidt Decomposition:

$$|AR\rangle = \sum_i \lambda_i|i_A\rangle|i_R\rangle$$

for orthonormal bases $|i_A\rangle, |i_R\rangle$ of A and R respectively. Define a linear transformation U such that

$\lambda_i|i_A\rangle = \sum_j U_{ij}p_j|\psi_j\rangle$. We then have that:

$$|AR\rangle = \sum_i \left(\sum_j U_{ij}p_j|\psi_j\rangle\right)|i_R\rangle = \sum_j p_j|\psi_j\rangle \sum_i U_{ij}|i_R\rangle.$$

We note that we can move the U_{ij} to system R as R has the same state space as A by construction. Letting $|j\rangle = \sum_i U_{ij}|i_R\rangle$ be our orthonormal basis of R, the claim follows (by part (2) of the question.

□

Problem 2.1: Functions of the Pauli matrices

Let $f(\cdot)$ be any function from complex numbers to complex numbers. Let $\mathbf{n}$ be a normalized vector in three dimensions, and let θ be real. Show

$$f(\theta\mathbf{n}\cdot\boldsymbol{\sigma}) = \frac{f(\theta)+f(-\theta)}{2}I + \frac{f(\theta)-f(-\theta)}{2}\mathbf{n}\cdot\boldsymbol{\sigma}$$

Solution

Concepts Involved: Linear Algebra, Spectral Decomposition, Operator Functions.

From Exercise 2.35, we recall that $\mathbf{n}\cdot\boldsymbol{\sigma}$ has spectral decomposition $\mathbf{n}\cdot\boldsymbol{\sigma} = |n_+\rangle\langle n_+| - |n_-\rangle\langle n_-|$. We then have that (by the definition of operator functions):

$$f(\theta\mathbf{n}\cdot\boldsymbol{\sigma}) = f\left(\theta(|n_+\rangle\langle n_+| - |n_-\rangle\langle n_-|)\right) = f(\theta)\,|n_+\rangle\langle n_+| + f(-\theta)\,|n_-\rangle\langle n_-|\,.$$

We then use the fact proven in the solution to Exercise 2.60 that we can write the projectors $P_\pm = |n_\pm\rangle\langle n_\pm|$ in terms of the operator $\mathbf{n}\cdot\boldsymbol{\sigma}$ as:

$$|n_\pm\rangle\langle n_\pm| = \frac{I \pm \mathbf{n}\cdot\boldsymbol{\sigma}}{2}.$$

Hence making this substitution we have:

$$f(\theta\mathbf{n}\cdot\boldsymbol{\sigma}) = f(\theta)\left(\frac{I+\mathbf{n}\cdot\boldsymbol{\sigma}}{2}\right) + f(-\theta)\left(\frac{I-\mathbf{n}\cdot\boldsymbol{\sigma}}{2}\right).$$

Grouping terms, we obtain the desired relation:

$$f(\theta\mathbf{n}\cdot\boldsymbol{\sigma}) = \frac{f(\theta)+f(-\theta)}{2}I + \frac{f(\theta)-f(-\theta)}{2}\mathbf{n}\cdot\boldsymbol{\sigma}.$$

□

Remark:
Arguably, the most used application of the above identity in quantum information is when $f(\theta\mathbf{n}\cdot\boldsymbol{\sigma}) =$

$\exp\{i(\theta/2)\mathbf{n}\cdot\boldsymbol{\sigma}\}$. In this case (as in Exercise 2.35), we have

$$\begin{aligned}\exp\{i(\theta/2)\mathbf{n}\cdot\boldsymbol{\sigma}\} &= \frac{\exp\{\theta/2\}+\exp\{-\theta/2\}}{2}I + \frac{\exp\{\theta/2\}-\exp\{-\theta/2\}}{2}\mathbf{n}\cdot\boldsymbol{\sigma} \\ &= \cos\left(\frac{\theta}{2}\right)I + i\sin\left(\frac{\theta}{2}\right)\mathbf{n}\cdot\boldsymbol{\sigma} \quad .\end{aligned}$$

Problem 2.2: Properties of Schmidt numbers

Suppose $|\psi\rangle$ is a pure state of a composite system with components A and B.

(1) Prove that the Schmidt number of $|\psi\rangle$ is equal to the rank of the reduced density matrix $\rho_A \equiv \mathrm{tr}_B(|\psi\rangle\langle\psi|)$. (Note that the rank of a Hermitian operator is equal to the dimension of its support.)

(2) Suppose $|\psi\rangle = \sum_j |\alpha_j\rangle|\beta_j\rangle$ is a representation for $|\psi\rangle$, where $|\alpha_j\rangle$ and $|\beta_j\rangle$ are (un-normalized) states for systems A and B, respectively. Prove that the number of terms in a such a decomposition is greater than or equal to the Schmidt number of $|\psi\rangle$, $\mathrm{Sch}(\psi)$.

(3) Suppose $|\psi\rangle = \alpha|\varphi\rangle + \beta|\gamma\rangle$. Prove that

$$\mathrm{Sch}(\psi) \geq \left|\mathrm{Sch}(\varphi) - \mathrm{Sch}(\gamma)\right|$$

Solution

Concepts Involved: Linear Algebra, Schmidt Decomposition, Schmidt Number, Reduced Density Operators.

---

(1) We write the Schmidt decomposed $|\psi\rangle$, and therefore the density matrix ρ_ψ as:

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle \implies |\psi\rangle\langle\psi| = \sum_{ii'} \lambda_i^2 |i_A\rangle\langle i_A| \otimes |i_B\rangle\langle i_B|$$

Taking the partial trace of subsystem B in the $|i_B\rangle$ basis, we obtain the reduced density matrix ρ_A to be:

$$\rho_A = \mathrm{tr}_B(|\psi\rangle\langle\psi|) = \sum_i \lambda_i^2 |i_A\rangle\langle i_A|$$

$\mathrm{Sch}(\psi)$ of the λ_is are nonzero, and therefore ρ_A has $\mathrm{Sch}(\psi)$ nonzero eigenvalues - therefore the rank of its support is $\mathrm{Sch}(\psi)$.

(2) Suppose for the sake of contradiction that some decomposition $|\psi\rangle = \sum_{j=1}^N |\alpha_j\rangle |\beta_j\rangle$ had less terms than the Schmidt decomposition of $|\psi\rangle$, i.e. $N <$ Sch(ψ).

The density matrix of $|\psi\rangle$ is:

$$\rho_\psi = |\psi\rangle\langle\psi| = \sum_{j=1,k=1}^{N} |\alpha_j\rangle\langle\alpha_k| \otimes |\beta_j\rangle\langle\beta_k| \tag{1}$$

Tracing out subsystem B, we obtain the reduced density matrix of subsystem A:

$$\rho_A = \mathrm{Tr}_B(\rho_\psi) = \sum_{j=1,k=1}^{N} |\alpha_j\rangle\langle\alpha_k| \langle\beta_j|\beta_k\rangle \tag{2}$$

where we have used that $\mathrm{Tr}\big(|\beta_1\rangle\langle\beta_2|\big) = \langle\beta_1|\beta_2\rangle$. From the above, it is clear that ρ_A has rank at most N, as the support of ρ_A is spanned by $\big\{|\alpha_1\rangle, \dots, |\alpha_N\rangle\big\}$. But then the rank of ρ_A is less than Sch(ψ), which contradicts our finding in part (a).

(3) If $\mathrm{Sch}(\varphi) = \mathrm{Sch}(\gamma)$ then there is nothing to prove as $\mathrm{Sch}(\psi)$ is non-negative by definition. Suppose then that $\mathrm{Sch}(\varphi) \neq \mathrm{Sch}(\gamma)$. WLOG suppose $\mathrm{Sch}(\varphi) > \mathrm{Sch}(\gamma)$. We can then write:

$$|\varphi\rangle = \frac{\beta}{\alpha}|\gamma\rangle - \frac{1}{\alpha}|\psi\rangle$$

If we Schmidt decompose $|\varphi\rangle$ and $|\psi\rangle$, we have written $|\varphi\rangle$ as the sum of $\mathrm{Sch}(\gamma) + \mathrm{Sch}(\psi)$ (unnormalized) bipartite states. Applying the result from part (2) of this problem, we then have that:

$$\mathrm{Sch}(\varphi) \leq \mathrm{Sch}(\gamma) + \mathrm{Sch}(\psi)$$

which we rearrange to obtain:

$$\mathrm{Sch}(\psi) \geq \mathrm{Sch}(\varphi) - \mathrm{Sch}(\gamma) = \big|\mathrm{Sch}(\varphi) - \mathrm{Sch}(\gamma)\big|$$

which proves the claim.

□

Problem 2.3: Tsirelson's inequality

Suppose $Q = \mathbf{q}\cdot\boldsymbol{\sigma}$, $R = \mathbf{r}\cdot\boldsymbol{\sigma}$, $S = \mathbf{s}\cdot\boldsymbol{\sigma}$, $T = \mathbf{t}\cdot\boldsymbol{\sigma}$, where $\mathbf{q}, \mathbf{r}, \mathbf{s}$, and $\mathbf{t}$ are real unit vectors in three dimensions. Show that

$$(Q\otimes S + R\otimes S + R\otimes T - Q\otimes T)^2 = 4I + [Q,R]\otimes[S,T]$$

Use this result to prove that

$$\langle Q\otimes S\rangle + \langle R\otimes S\rangle + \langle R\otimes T\rangle - \langle Q\otimes T\rangle \leq 2\sqrt{2}$$

so the violation of the Bell inequality found in Equation (2.230) is the maximum possible in quantum mechanics.

Solution

Concepts Involved: Tensor Products, Commutators

We first show that $N^2 = I$ for any $N = \mathbf{n} \cdot \boldsymbol{\sigma}$ where $\mathbf{n}$ is a unit vector in three dimensions. We have that:

$$\begin{aligned} N^2 &= (\sum_{i=1}^{3} n_i \sigma_i)^2 \\ &= n_1^2\sigma_1^2 + n_2^2\sigma_2^2 + n_3^2\sigma_3^2 + n_1 n_2(\sigma_1\sigma_2 + \sigma_2\sigma_1) + n_1 n3(\sigma_1\sigma_3 + \sigma_3\sigma_1) + n_2 n_3(\sigma_2\sigma_3 + \sigma_3\sigma_2) \end{aligned}$$

By Exercise 2.41, $\sigma_i^2 = I$ and $\left\{\sigma_i, \sigma_j\right\} = 0$ for $i \neq j$, so the above reduces to:

$$N^2 = n_1^2 I + n_2^2 I + n_3^2 I = (n_1^2 + n_2^2 + n_3^2)I = I$$

where we use the fact that $\mathbf{n}$ is of unit length. Using this fact, we have that:

$$\begin{aligned} (Q \otimes S + R \otimes S + R \otimes T - Q \otimes T)^2 &= Q^2 \otimes S^2 + QR \otimes S^2 + QR \otimes ST - Q^2 \otimes ST \\ &+ RQ \otimes S^2 + R^2 \otimes S^2 + R^2 \otimes ST - RQ \otimes ST \\ &+ RQ \otimes TS + R^2 \otimes TS + R^2 \otimes T^2 - RQ \otimes T^2 \\ &- Q^2 \otimes TS - QR \otimes TS - QR \otimes T^2 + Q^2 \otimes T^2 \\ &= I \otimes I + QR \otimes I + QR \otimes ST - I \otimes ST \\ &+ RQ \otimes I + I \otimes I + I \otimes ST - RQ \otimes ST \\ &+ RQ \otimes TS + I \otimes TS + I \otimes I - RQ \otimes I \\ &- I \otimes TS - QR \otimes TS - QR \otimes I + I \otimes I \\ &= 4I \otimes I + RQ \otimes TS - RQ \otimes ST + QR \otimes ST - QR \otimes TS \\ &= 4I + QR \otimes (ST - TS) - RQ \otimes (ST - TS) \\ &= 4I + [Q, R] \otimes [S, T] \end{aligned}$$

which proves the first equation. We have that $\langle 4I \rangle = 4 \langle I \rangle = 4$. Since each of Q, R, S, T have eigenvalues ± 1 (Exercise 2.35), we also ave that $\left\langle [Q, R] \otimes [S, T] \right\rangle \leq 4$ as the tensor product of commutators consists of 4 terms, each of which has expectation less than or equal to 1. We therefore have by the linearity of expectation (Exercise A1.4) that:

$$\left\langle (Q \otimes S + R \otimes S + R \otimes T - Q \otimes T)^2 \right\rangle = \left\langle 4I + [QR] \otimes [S, T] \right\rangle \leq 8.$$

Furthermore, we have that:

$$\left\langle (Q \otimes S + R \otimes S + R \otimes T - Q \otimes T) \right\rangle^2 \leq \left\langle (Q \otimes S + R \otimes S + R \otimes T - Q \otimes T)^2 \right\rangle$$

so combining the two inequalities we obtain:

$$\left\langle (Q \otimes S + R \otimes S + R \otimes T - Q \otimes T) \right\rangle^2 \leq 8.$$

Taking square roots on both sides, we have:

$$\left\langle (Q \otimes S + R \otimes S + R \otimes T - Q \otimes T) \right\rangle \leq 2\sqrt{2}$$

and again by the linearity of expectation:

$$\langle Q \otimes S \rangle + \langle R \otimes S \rangle + \langle R \otimes T \rangle - \langle Q \otimes T \rangle \leq 2\sqrt{2}$$

which is the desired inequality. $\square$

3 Introduction to computer science

Exercise 3.1: Non-computable processes in Nature

How might we recognize that a process in Nature computes a function not computable by a Turing machine?

Solution

Concepts Involved: Turing Machines.

One criteria is natural phenomena that appear to be truly random; Turing machines as defined in the text are deterministic (though there are probabilistic variations that would solve this issue) and hence would not be able to compute a random function. From a more direct point, if a process in Nature was to be found to compute a known non-computable problem (e.g. solve the Halting problem or the Tiling problem) then we may conclude (trivially) that the process would not be computable. However since the domain of inputs that we could provide top such a natural process would have to be finite, there would be no concrete method in which one could actually test if such a process was truly computing a non-Turing computable function (as a Turing machine that works on a finite subset of inputs for an uncomputable problem could be devised). □

Exercise 3.2: Turing numbers

Show that single-tape Turing machines can each be given a number from the list $1, 2, 3, \ldots$ in such a way that the number uniquely specifies the corresponding machine. We call this number the *Turing number* of the corresponding Turing machine. (Hint: Every positive integer has a unique prime factorization $p_1^{a_1} p_2^{a_2} \ldots p_k^{a_k}$, where p_i are distinct prime numbers, and $a_1, \ldots a_k$ are non-negative integers.)

Exercise 3.3: Turing machine to reverse a bit string

Describe a Turing machine which takes a binary number x as input, and outputs the bits of x in reverse order. (*Hint:* In this exercise and the next it may help to use a multi-tape Turing machine and/or symbols other than $\triangleright, 0, 1$ and the blank.)

Exercise 3.4: Turing machine to add modulo 2

Describe a Turing machine to add two binary numbers x and y modulo 2. The numbers are input on the Turing machine tape in binary, in the form x, followed by a single blank, followed by a y. If one number is not as long as the other then you may assume that it has been padded with leading 0s to make the two numbers the same length.

Exercise 3.5: Halting problem with no inputs

Show that given a Turing machine M there is no algorithm to determine whether M halts when the input to the machine is a blank tape.

Exercise 3.6: Probabilistic halting problem

Suppose we number the probabilistic Turing machines using a scheme similar to that found in Exercise 3.2 and define the probabilistic halting function $h_p(x)$ to be 1 if machine x halts on input of x with probability at least 1/2 and 0 if machine x halts on input of x with probability less than 1/2. Show that there is no probabilistic Turing machine which can output $h_p(x)$ with probability of correctness strictly greater than 1/2 for all x.

Exercise 3.7: Halting oracle

Suppose a *black box* is made available to us which takes a non-negative integer x as input, and then outputs the value of $h(x)$, where $h(\cdot)$ is the halting function defined in Box 3.2 on page 130. This type of black box is sometimes known as an *oracle* for the halting problem. Suppose we have a regular Turing machine which is augmented by the power to call the oracle. One way of accomplishing this is to use a two-tape Turing machine, and add an extra program instruction to the Turing machine which results in the oracle being called, and the value of $h(x)$ being printed on the second tape, where x is the current contents of the second tape. It is clear that this model for computation is more powerful than the conventional Turing machine model, since it can be used to compute the halting function. Is the halting problem for this model of computation undecidable? That is, can a Turing machine aided by an oracle for the halting problem decide whether a program for the Turing machine with oracle will halt on a particular input?

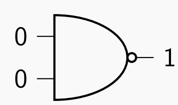
Exercise 3.8: Universality of `NAND`

Show that the `NAND` gate can be used to simulate the `AND`, `XOR`, and `NOT` gates, provides wires, ancilla bits and `FANOUT` are available.

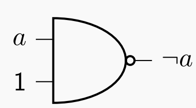
Solution

Concepts Involved: Logic Gates.

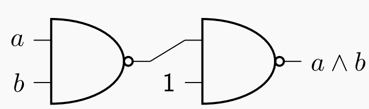
We start by showing how we can get a 1 qubit using two 0 ancilla bits and a `NAND` gate.



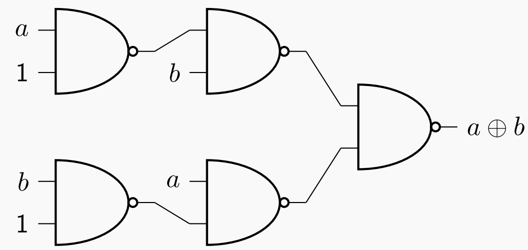
We will now show how to simulate the `NOT`, `AND`, and `XOR` gates. We note that we will use "1" to denote as shorthand a 1 bit constructed using two ancilla bits (as above). a/b represent the input bits. We start with the `NOT` gate.



Next, we simulate the `AND` gate.



For the `XOR` simulated gate, we note that we first use `FANOUT` twice to copy both input bits.



Having simulated the three gates using the `NAND` gate only, we conclude that the `NAND` is universal. □

Exercise 3.9

Prove that $f(n)$ is $O(g(n))$ if and only if $g(n)$ is $\Omega(f(n))$. Deduce that $f(n)$ is $\Theta(g(n))$ if and only if $g(n)$ is $\Theta(f(n))$.

Solution

Concepts Involved: Asymptotic Notation.

Suppose $f(n)$ is $O(g(n))$. Then, there exists $c > 0$ such that for all $n > n_0$, $f(n) \leq cg(n)$. Therefore, we have $\frac{1}{c} > 0$ such that for all $n > n_0$, $\frac{1}{c}f(n) \leq g(n)$. Hence, $g(n)$ is $\Omega(f(n))$. Conversely, if $g(n)$ is $\Omega(f(n))$, there exists $c > 0$ such that for all $n > n_0$, $cf(n) \leq g(n)$. Hence, we have $\frac{1}{c} > 0$ such that for all $n > n_0$, $f(n) \leq \frac{1}{c}g(n)$ and hence $f(n)$ is $O(g(n))$.
Therefore, if $f(n)$ is $\Theta(g(n))$ then $f(n)$ is $O(g(n))$ and $\Omega(g(n))$, and by the above argument, $g(n)$ is $O(f(n))$ and $\Omega(f(n))$ and hence $g(n)$ is $\Theta(f(n))$. The converse holds in the same way. □

Exercise 3.10

Suppose $g(n)$ is a polynomial of degree k. Show that $g(n)$ is $O(n^l)$ for any $l \geq k$.

Solution

Concepts Involved: Asymptotic Notation.

By assumption, $g(n) = a_0 + a_1 n^1 + a_2 n^2 + \ldots + a_k n^k$ with $a_k \neq 0$. For $n \geq 1$ we have that $n^l \geq n^k$ if $l \geq k$, and hence if $l \geq k$ we have that $a_i n^l \geq a_i n^i$ for all $i \in \{0, \ldots, k\}$. Therefore, we have that:

$$(a_0 + a_1 + \ldots + a_k)n^l \geq a_0 + a_1 n^1 + \ldots + a_k n^k = g(n)$$

for $n \geq 1$ and hence $g(n)$ is $O(n^l)$. □

Exercise 3.11

Show that $\log n$ is $O(n^k)$ for any $k > 0$.

Solution

Concepts Involved: Asymptotic Notation.

Let $k > 0$ and $c > 0$. By the definition of the exponential we have that:

$$\exp\left(cn^k\right) = \sum_{j=0}^{\infty} \frac{(an^k)^j}{j!} = \sum_{j=0}^{\infty} \frac{a^{kj} n^{kj}}{j!}$$

now, there exists some $j_0 \in \mathbb{Z}$ for which $kj_0 > 1$. Since for $n \geq 0$ the terms in the above sum are non-negative, we find:

$$\exp\left(cn^k\right) \geq \frac{c^{kj_0} n^{kj_0}}{j_0!}$$

Now, choose c sufficiently large such that $c^{kj_0} \geq j_0!$. We then find that:

$$\exp\left(cn^k\right) \geq \frac{c^{kj_0} n^{kj_0}}{j_0!} \geq n^{kj_0}$$

Then for $n \geq 1$ it follows that $n^{kj_0} \geq n$ as $kj_0 \geq 1$ and so:

$$\exp\left(cn^k\right) \geq n$$

Since the logarithm is monotonic, we may take the log of both sides and preserve the inequality:

$$cn^k \geq \log n$$

So we have shown that for any $k > 0$, there exists $c > 0$ such that for all $n > 1$, $cn^k \geq \log n$. Hence, $\log n$ is $O(n^k)$ for any $k > 0$. $\square$

Exercise 3.12: $n^{\log n}$ is super-polynomial

Show that n^k is $O(n^{logn})$ for any k, but that $n^{\log n}$ is never $O(n^k)$.

Solution

Concepts Involved: Asymptotic Notation.

First, note that for any k, $e^k \leq n$ for sufficiently large $n > n_0$ and so $k \leq \log n$ by monotonicity of the logarithm. Therefore, for $n > n_0$ it follows by monotonicity (of exponentiation) that $n^k \leq n^{\log n}$ and so n^k is $O(n^{\log n})$.
Now, consider an arbitrary $a > 0$. It still follows for sufficiently large $n > n_0$ that $e^{ak} \leq n$ and so $ak \leq \log n$ and $n^a n^k \leq n^{\log n}$. But for any $c > 0$ $n^a > c$ for sufficiently large n and so:

$$cn^k \leq n^{\log n}$$

So since for any $c > 0$ there exists some n_0' for which $n > n_0'$ implies $cn^k \leq n^{\log n}$, it follows that $n^{\log n}$ is never $O(n^k)$. $\square$

Exercise 3.13: $n^{\log n}$ is sub-exponential

Show that c^n is $\Omega(n^{\log n})$ for any $c > 1$, but that $n^{\log n}$ is never $\Omega(c^n)$.

Solution

Concepts Involved: Asymptotic Notation.

First note from Exercise 1.11 that $\log n$ is $O(n^k)$ for any $k > 0$. Specifically, take $k = 1/2$; then there exists $a > 0$ such that for $n > n_0$:

$$an^{1/2} \geq \log n$$

and therefore squaring both sides:

$$a^2 n \geq \log n \log n = \log n^{\log n}$$

Now for any $c > 1$, we can define $a' = \frac{a^2}{\log c} > 0$ and write:

$$na' \log c = \log c^{na'} \geq \log n^{\log n}$$

Exponentiating both sides preserves the inequality, and so:

$$c^{na'} = c^{a'} c^n \geq n^{\log n}$$

and so there exists a constant $\frac{1}{c^{a'}} > 0$ such that for $n > n_0$, $c^n \geq \frac{1}{c^{a'}} n^{\log n}$ and therefore c^n is $\Omega(n^{\log n})$. Now, let $b > 0$ be some arbitrarily small constant. For sufficiently large n, we have that $na' logc + \log b > 0$ and so for sufficiently large n it further follows that:

$$\log b + na' \log c = \log\left(bc^{na'}\right) \geq \log n^{\log n}$$

where a' is defined as it was previously. Therefore exponentiating both sides:

$$bc^{na'} = bc^{a'} c^n \geq n^{\log n}$$

so for sufficiently large n, for any arbitrarily small constant b' it follows that $b' c^n \geq n^{\log n}$ and so $n^{\log n}$ is never $\Omega(c^n)$. $\square$

Exercise 3.14

Suppose $e(n)$ is $O(f(n))$ and $g(n)$ is $O(h(n))$. Show that $e(n)g(n)$ is $O(f(n)h(n))$.

Solution

Concepts Involved: Asymptotic Notation.

By assumption, we have that $e(n) \leq c_1 f(n)$ for some $c_1 > 0$ and for all $n > n_1$ and that $g(n) \leq c_2 h(n)$

for some $c_2 > 0$ and for all $n > n_2$. Let $n_0 = \max n_1, n_2$. We then have that for $n > n_0$ that:

$$e(n)g(n) \leq c_1 f(n) c_2 h(n) = (c_1 c_2)(f(n)h(n))$$

so therefore $e(n)g(n)$ is $O(f(n)h(n))$. □

Exercise 3.15: Lower bound for compare-and-swap based sorts

Suppose an n element list is sorted by applying some sequence of compare-and-swap operations to the list. There are $n!$ possible initial orderings of the list. Show that after k of the compare-and-swap operations have been applied, at most 2^k of the possible initial orderings will have been sorted into the correct order. Conclude that $\Omega(n \log n)$ compare and swap operations are required to sort all possible initial orderings into the correct order.

Solution

Concepts Involved: Asymptotic Notation, Compare-and-Swap.

We prove the first statement by induction. After 0 steps, we have that $1 = 2^0$ out of the $n!$ possible orderings are already sorted. Let $k \in \mathbb{N}, k \geq 0$ and suppose that after k swaps, at most 2^k of the initial orderings have been sorted into the correct order. We now consider the state of the list after the $k+1$th swap. Each of the 2^k initial orderings from the previous step are correctly sorted already (so the swap does nothing), and there are a further 2^k initial orderings that are one swap away from the 2^k from the previous step, and hence the $k+1$th swap will put 2^k more initial orderings into the correct order. Therefore, after 2^{k+1} compare and swaps, there are at most $2^k + 2^k = 2^{k+1}$ possible initial orderings that are sorted into the correct order. This proves the claim.

Using the above fact, we have that in order to have all $n!$ possible initial orderings correct after k steps that $2^k \geq n!$. Taking logarithms on both sides, we have that $\log(2^k) \geq \log(n!)$ and hence $k \geq \log(n!)$. Using Stirling's approximation for factorials (`https://en.wikipedia.org/wiki/Stirling%27s_approximation`), we have that:

$$k \geq n \log n - n \log e + O(\log n)$$

from which we conclude that k is $\Omega(n \log n)$ and hence $\Omega(n \log n)$ compare and swap operations are required to sort all possible initial orderings into the correct order. □

Exercise 3.16: Hard-to-compute functions exist

Show there exist Boolean functions on n inputs which require at least $2^n / \log n$ logic gates to compute.

Exercise 3.17

Prove that a polynomial-time algorithm for finding the factors of a number m exists if and ony if the factoring decision problem is in $\mathbf{P}$.

Exercise 3.18

Prove that if $\mathbf{coNP} \neq \mathbf{NP}$ then $\mathbf{P} \neq \mathbf{NP}$.

Exercise 3.19

The `Reachability` problem is to determine whether there is a path between two specified vertices in a graph. Show that `Reachability` can be solved using $O(n)$ operations if the graph has n vertices. Use the solution to `Reachability` to show that it is possible to decide whether a graph is connected in $O(n^2)$ operations.

Exercise 3.20: Euler's theorem

Prove Euler's theorem. In particular, if each vertex has an even number of incident edges, give a constructive procedure for finding an Euler cycle.

Exercise 3.21: Transitive property of reduction

Show that if a language L_1 is reducible to the language L_2 and the language L_2 is reducible to L_3 then the language L_1 is reducible to the language L_3.

Exercise 3.22

Suppose L is complete for a complexity class, and L' is another language in the complexity class such that L reduces to L'. Show that L' is complete for the complexity class.

Exercise 3.23

Show that `SAT` is $\mathbf{NP}$-complete by first showing that the `SAT` is in $\mathbf{NP}$, and then showing that `CSAT` reduces to `SAT`.

Exercise 3.24: `2SAT` has an efficient solution

Suppose φ is a Boolean formula in conjunctive normal form, in which each clause contains only two literals.

(1) Construct a (directed) graph $G(\varphi)$ with directed edges in the following way: the vertices of G correspond to variables x_k and their negations $\neg x_j$ in φ. There is a (directed) edge (α, β) in G if and only if the clause $(\neg\alpha \vee \beta)$ or the clause $(\beta \wedge \neg\alpha)$ is present in φ. Show that φ is not satisfiable if and only if there exists a variable x such that there are paths from x and $\neg x$ and from $\neg x$ to x in $G(\varphi)$.

(2) Show that given a directed graph G containing n vertices it is possible to determine whether two vertices v_1 and v_2 are connected in polynomial time.

(3) Find an efficient algorithm to solve `2SAT`.

Exercise 3.25: $\mathbf{PSPACE} \subseteq \mathbf{EXP}$

The complexity class $\mathbf{EXP}$ (for *exponential time*) contains all decision problems which may be decided by a Turing machine running in exponential time, that is time $O(2^{n^k})$, where k is any constant. Prove that $\mathbf{PSPACE} \subseteq \mathbf{EXP}$. (*Hint:* If a Turing machine has l internal states, an m letter alphabet, and uses space $p(n)$, argue that the machine can exist in one of at most $lm^{p(n)}$ different states, and that if the Turing machine is to avoid infinite loops then it must halt before revisiting a state.)

Exercise 3.26: $\mathbf{L} \subseteq \mathbf{P}$

The complexity class $\mathbf{L}$ (for *logarithmic space*) contains all decision problems which may be decided by a Turing machine running in logarithmic space, that is, in space $O(\log(n))$. More precisely, the class $\mathbf{L}$ is defined using a two-tape Turing machine. The first tape contains the problem instance, of size n, and is a read-only tape, in the sense that only program lines which don't change the contents of the first tape are allowed. The second tape is a working tape which initially contains only blanks. The logarithmic space requirement is imposed on the second, working tape only. Show that $\mathbf{L} \subseteq \mathbf{P}$.

Exercise 3.27: Approximation algorithm for `VERTEX COVER`

Let $G = (V, E)$ be an undirected graph. Prove that the following algorithm finds a vertex cover for G that is within a factor of two of being a minimal vertex cover.

$VC = \emptyset$
$E' = E$
while $E' = \emptyset$ **do**
  let (α, β) be any edge of E'
  $VC = VC \cup \{\alpha, \beta\}$
  remove from E' every edge incident on α or β
end
return *VC*

Exercise 3.28: Arbitrariness of the constant in the definition of BPP

Suppose k is a fixed constant, $1/2 < k \leq 1$. Suppose L is a language such that there exists a Turing machine M with the property that whenever $x \in L$, M accepts x with probability at least k, and whenever $x \notin L$, M rejects x with probability at least k. Show that $L \in \mathbf{BPP}$.

Exercise 3.29: Fredkin gate is self-inverse

Show that applying two consecutive Fredkin gates gives the same outputs as inputs.

Solution

Concepts Involved: Fredkin Gates. Recall the input/output table of the Fredkin gate:

Inputs			Outputs		
a	b	c	a'	b'	c'
0	0	0	0	0	0
0	0	1	0	0	1
0	1	0	0	1	0
0	1	1	1	0	1
1	0	0	1	0	0
1	0	1	0	1	1
1	1	0	1	1	0
1	1	1	1	1	1

We check for all possible 8 input states that applying the Fredkin gate returns the original input state.

$$\begin{aligned}
F[F[(0,0,0)]] &= F[(0,0,0)] = (0,0,0) \\
F[F[(0,0,1]] &= F[(0,0,1)] = (0,0,1) \\
F[F[(0,1,0]] &= F[(0,1,0)] = (0,1,0) \\
F[F[(0,1,1)]] &= F[(1,0,1)] = (0,1,1) \\
F[F[(1,0,0)]] &= F[(1,0,0)] = (1,0,0) \\
F[F[(1,0,1)]] &= F[(0,1,1)] = (1,0,1) \\
F[F[(1,1,0)]] &= F[(1,1,0)] = (1,1,0) \\
F[F[(1,1,1)]] &= F[(1,1,1)] = (1,1,1)
\end{aligned}$$

We conclude that the Fredkin gate is self-inverse. □

Exercise 3.30

Verify that the billiard ball computer in Figure 3.14 computes the Fredkin gate.

Exercise 3.31: Reversible half-adder

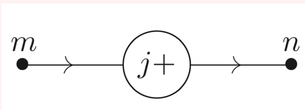
Construct a reversible circuit which, when two bits x and y are input, outputs $(x, y, c, x \oplus y)$, where c is the carry bit when x and y are odd.

Exercise 3.32: From Fredkin to Toffoli and back again

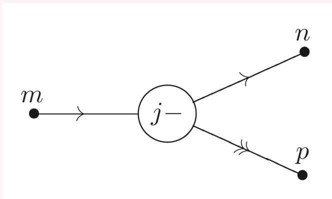
What is the smallest number of Fredkin gates needed to simulate a Toffoli gate? What is the smallest number of Toffoli gates needed to simulate a Fredkin gate?

Problem 3.1: Minsky machines

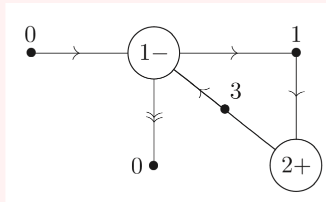
A *Minsky machine* consists of a finite set of *registers*, $r_1, r_2, \ldots, r_k$, each capable of holding an arbitrary non-negative integer, and a *program*, made up of *orders* of one of two types. The first type has the form:



The interpretation is that at point m in the program register rj is incremented by one, and execution proceeds to point n in the program. The second type of order has the form:



The interpretation is that at point m in the program, register r_j is decremented if it contains a positive integer, and execution proceeds to point n in the program. If register r_j is zero then execution simply proceeds to point p in the program. The *program* for the Minsky machine consists of a collection of such orders, of a form like:



The starting and all possible halting points for the program are conventionally labeled zero. This program takes the contents of register r_1 and adds them to register r_2, while decrementing r_1 to zero.

(1) Prove that all (Turing) computable functions can be computed on a Minsky machine, in the sense that given a computable function $f(\cdot)$ there is a Minsky machine program that when the registers start in the state $(n, 0, \dots, 0)$ gives as output $(f(n), 0, \dots, 0)$.

(2) Sketch a proof that any function which can be computed on a Minsky machine, in the sense just defined, can also be computed on a Turing machine.

Problem 3.2: Vector games

A *vector game* is specified by a finite list of vectors, all of the same dimension, and with integer co-ordinates. The game is to start with a vector x of non-negative integer co-ordinates and to add to x the first vector from the list which preserves the non-negativity of all the components, and to repeat this process until it is no longer possible. Prove that for any computable function $f(\cdot)$ there is a vector game which when started with the vector $(n, 0, \dots, 0)$ reaches $(f(n), 0, \dots, 0)$ (*Hint:* Show that a vector game in $k + 2$ dimensions can simulate a Minsky machine containing k registers.)

Problem 3.3: Fractran

A *Fractran* program is defined by a list of positive rational numbers $q_1, \dots, q_n$. It acts on a positve integer m by replacing it by $q_i m$ where i is the least number such that $q_i m$ is an integer. If there is ever a time when there is no i such that $q_i m$ is an integer, then execution stops. Prove that for any computable function $f(\cdot)$ there is a Fractran program which when started with 2^n reaches $2^{f(n)}$ without going through any intermediate powers of 2. (*Hint:* use the previous problem.)

Problem 3.4: Undecidability of dynamical systems

A Fractran program is essentially just a very simple dynamical system taking positive integers to positive integers. Prove that there is no algorithm to decide whether such a dynamical system ever reaches 1.

Problem 3.5: Non-universality of two bit reversible logic

Suppose we are trying to build circuits using only one and two bit reversible logic gates, and ancilla bits. Prove that there are Boolean functions which cannot be computed in this fashion. Deduce that the Toffoli gate cannot be simulated using one and two bit reversible gates, even with the aid of ancilla bits.

Problem 3.6: Hardness of approximation of TSP

Let $r \geq 1$ and suppose that there is an approximation algorithm for TSP which is guaranteed to find the shorted tour among n cities to within a factor r. Let $G = (V, E)$ be any graph on n vertices. Define an instance of TSP by identifying cities with vertices in V, and deefinign the distance between cities i and j to be 1 if (i, j) is an edge of G, and to be $\lceil r \rceil |V| + 1$ otherwise. Show that if the approximation algorithm is applied to this instance of TSP then it returns a Hamiltonian cycle for G if one exists, and otherwise returns a tour of length more than $\lceil r \rceil |V|$. From the $\mathbf{NP}$-completeness of HC it follows that no such approximation algorithm can exist unless $\mathbf{P} = \mathbf{NP}$.

Problem 3.7: Reversible Turing machines

(1) Explain how to construct a reversible Turing machine that can compute the same class of functions as is computable on an ordinary Turing machine. (*Hint:* It may be helpful to use a multi-tape construction.)

(2) Give general space and time bounds for the operation of your reversible Turing machine, in terms of the time $t(x)$ and space $s(x)$ required on an ordinary single-tape Turing machine to compute a function $f(x)$.

Problem 3.8: Find a hard-to-compute class of functions (Research)

Find a natural class of functions on n inputs which requires a super-polynomial number of Boolean gates to compute.

Problem 3.9: Reversible $\mathbf{PSPACE} = \mathbf{PSPACE}$

It can be shown that the problem 'quantified satisfiability', or QSAT, is $\mathbf{PSPACE}$-complete. That is, every other language in $\mathbf{PSPACE}$ can be reduced to QSAT in polynomial time. The language QSAT is defined to consist of all Boolean formulae φ in n variables $x_1, \ldots, x_n$, and in conjunctive normal form, such that:

$$\exists_{x_1} \forall_{x_2} \exists_{x_3} \ldots \forall_{x_n} \varphi \text{ if } n \text{ is even;}$$
$$\exists_{x_1} \forall_{x_2} \exists_{x_3} \ldots \exists_{x_n} \varphi \text{ if } n \text{ is odd.}$$

Prove that a reversible Turing machine operating in polynomial space can be used to solve QSAT. Thus, the class of languages decidable by a computer operating reversibly in polynomial space is equal to $\mathbf{PSPACE}$.

Problem 3.10: Ancilla bits and efficiency of reversible computation

Let p_m be the mth prime number. Outline the construction of a reversible circuit which, upon the input of m and n such that $n > m$, outputs the product $p_m p_n$, that is $(m, n) \mapsto (p_m p_n, g(m, n))$ where $g(m, n)$ is the final state of the ancilla bits used by the circuit. Estimate the number of ancilla qubits your circuit requires. Prove that if a polynomial (in $\log n$) size reversible circuit can be found that uses $O(\log(\log n))$ ancilla bits then the problem of factoring a product of two prime numbers is in $\mathbf{P}$.

4 Quantum circuits

Exercise 4.1

In Exercise 2.11, which you should do now if you haven't already done it, you computed the eigenvectors of the Pauli matrices. Find the points on the Bloch sphere which correspond to the normalized eigenvectors of the different Pauli matrices.

Solution

Concepts Involved: Linear Algebra.
Recall that a single qubit in the state $|\psi\rangle = a|0\rangle + b|1\rangle$ can be visualized as a point (θ, φ) on the Bloch sphere, where $a = \cos(\theta/2)$ and $b = e^{i\varphi}\sin(\theta/2)$.

We recall from 2.11 that Z (and I) has eigenvectors $|0\rangle, |1\rangle$, X has eigenvectors $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}, |-\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}}$, and Y has eigenvectors $|y_+\rangle = \frac{|0\rangle+i|1\rangle}{\sqrt{2}}, |y_-\rangle = \frac{|0\rangle-i|1\rangle}{\sqrt{2}}$. Expressing these vectors as points on the Bloch sphere (using spherical coordinates), we have:

$$|0\rangle \cong (0,0)\,; |1\rangle \cong (\pi, 0)\,; |+\rangle \cong \left(\frac{\pi}{2}, 0\right);$$
$$|-\rangle \cong \left(\frac{\pi}{2}, \pi\right); |y_+\rangle \cong \left(\frac{\pi}{2}, \frac{\pi}{2}\right); |y_-\rangle \cong \left(\frac{\pi}{2}, \frac{3\pi}{2}\right).$$

□

Exercise 4.2

Let x be a real number and A a matrix such that $A^2 = I$. Show that

$$\exp(iAx) = \cos(x)I + i\sin(x)A$$

Use this result to verify Equations (4.4) through (4.6).

Solution

Concepts Involved: Linear Algebra, Operator Functions.

Let $|v\rangle$ be an eigenvector of A with eigenvalue λ. It then follows that $A^2|v\rangle = \lambda^2|v\rangle$, and furthermore we have that $A^2|v\rangle = I|v\rangle = |v\rangle$ by assumption. We obtain that $\lambda^2 = 1$ and therefore the only possible eigenvalues of A are $\lambda = \pm 1$. Let $|v_1\rangle, \ldots, |v_k\rangle$ be the eigenvectors with eigenvalue 1 and $|v_{k+1}\rangle, \ldots, |v_n\rangle$ be the eigenvectors with eigenvalue -1. By the spectral decomposition, we can write:

$$A = \sum_{i=1}^{k} |v_i\rangle\langle v_i| - \sum_{i=k+1}^{n} |v_i\rangle\langle v_i|$$

so by the definition of operator functions we have:

$$\exp(iAx) = \sum_{i=1}^{k} \exp(ix) \left|v_i\right\rangle\!\left\langle v_i\right| + \sum_{i=k+1}^{n} \exp(-ix) \left|v_i\right\rangle\!\left\langle v_i\right|.$$

By Euler's identity we have:

$$\exp(iAx) = \sum_{i=1}^{k} \big(\cos(x) + i\sin(x)\big) \left|v_i\right\rangle\!\left\langle v_i\right| + \sum_{i=k+1}^{n} \big(\cos(x) - i\sin(x)\big) \left|v_i\right\rangle\!\left\langle v_i\right|.$$

Grouping terms, we obtain:

$$\exp(iAx) = \cos(x)\sum_{i=1}^{n} \left|v_i\right\rangle\!\left\langle v_i\right| + i\sin(x)\left(\sum_{i=1}^{k} \left|v_i\right\rangle\!\left\langle v_i\right| - \sum_{i=k+1}^{n} \left|v_i\right\rangle\!\left\langle v_i\right|\right).$$

Using the spectral decomposition and definition of I, we therefore obtain the desired relation:

$$\exp(iAx) = \cos(x)I + i\sin(x)A.$$

Since all of the Pauli matrices satisfy $A^2 = I$ (Exercise 2.41), for $\theta \in \mathbb{R}$ we can apply this obtained relation to obtain:

$$\exp\big(-i\theta X/2\big) = \cos\left(\frac{\theta}{2}\right)I - i\sin\left(\frac{\theta}{2}\right)X = \begin{bmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}$$

$$\exp\big(-i\theta Y/2\big) = \cos\left(\frac{\theta}{2}\right)I - i\sin\left(\frac{\theta}{2}\right)Y = \begin{bmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}$$

$$\exp\big(-i\theta Z/2\big) = \cos\left(\frac{\theta}{2}\right)I - i\sin\left(\frac{\theta}{2}\right)Z = \begin{bmatrix} \cos\frac{\theta}{2} - i\sin\frac{\theta}{2} & 0 \\ 0 & \cos\frac{\theta}{2} - i\sin\frac{\theta}{2} \end{bmatrix} = \begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix}$$

which verifies equations (4.4)-(4.6). □

Exercise 4.3

Show that, up to a global phase, the $\pi/8$ gate satisfies $T = R_z(\pi/4)$

Solution

Concepts Involved: Linear Algebra, Quantum Gates.
Recall that the T gate is defined as:

$$T = \begin{bmatrix} 1 & 0 \\ 0 & \exp\big(i\pi/4\big) \end{bmatrix}$$

We observe that:

$$R_z(\pi/4) = \begin{bmatrix} e^{-i\pi/8} & 0 \\ 0 & e^{i\pi/8} \end{bmatrix} = e^{-i\pi/8} \begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix} = e^{-i\pi/8} T.$$

□

Exercise 4.4

Express the Hadamard gate H as a product of R_x and R_z rotations and $e^{i\varphi}$ for some φ.

Solution

Concepts Involved: Linear algebra, Quantum Gates

We claim that $H = R_z(\pi/2)R_x(\pi/2)R_z(\pi/2)$ up to a global phase of $e^{-i\pi/2}$. Doing a computation to verify this claim, we see that:

$$\begin{aligned}
R_z(\pi/2)R_x(\pi/2)R_z(\pi/2) &= \begin{bmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{bmatrix} \begin{bmatrix} \cos\frac{\pi}{4} & -i\sin\frac{\pi}{4} \\ -i\sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{bmatrix} \begin{bmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{bmatrix} \\
&= \begin{bmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{bmatrix} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{bmatrix} \begin{bmatrix} e^{-i\pi/4} & -ie^{i\pi/4} \\ -ie^{-i\pi/4} & e^{i\pi/4} \end{bmatrix} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} e^{-i\pi/2} & -i \\ -i & e^{i\pi/2} \end{bmatrix} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} e^{-i\pi/2} & -e^{-i\pi/2} \\ -e^{-i\pi/2} & e^{i\pi/2} \end{bmatrix} \\
&= \frac{e^{-i\pi/2}}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\
&= e^{-i\pi/2} H
\end{aligned}$$

□

Remark: If you are more algebraically minded, the following may appeal to you.

$$\begin{aligned}
R_z(\pi/2)R_x(\pi/2)R_z(\pi/2) &= \frac{1}{2\sqrt{2}}(1-iZ)(1-iX)(1-iZ) \\
&= \frac{1}{2\sqrt{2}}(1-iZ-iX-ZX)(1-iZ) \\
&= \frac{1}{2\sqrt{2}}(1-iZ-iX-ZX-iZ-XZ-1+iZXZ) \\
&= \frac{1}{2\sqrt{2}}(-2iX-2iZ) \quad (\text{using } ZXZ=-X) \\
&=: -iH
\end{aligned}$$

Exercise 4.5

Prove that $(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})^2 = I$, and use this to verify Equation (4.8)

Solution

Concepts Involved: Linear Algebra

Expanding out the expression, we see that:

$$\begin{aligned}
(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})^2 &= (n_xX+n_yY+n_zZ)^2 \\
&= n_x^2X^2+n_y^2Y^2+n_z^2Z^2+n_xn_y(XY+YX)+n_xn_z(XZ+ZX)+n_yn_z(YZ+ZY)
\end{aligned}$$

Using the result from Exercise 2.41 that $\{\sigma_i,\sigma_j\}=0$ if $i\neq j$ and $\sigma_i^2=I$, we have that:

$$(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})^2 = (n_x^2+n_y^2+n_z^2)I = I$$

where we use the fact that $\hat{\mathbf{n}}$ is a vector of unit length. With this shown, we can use the result of Exercise 4.2 to conclude that:

$$\exp(-i\theta\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}/2) = \cos\left(\frac{\theta}{2}\right) - i\sin\left(\frac{\theta}{2}\right)(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})$$

which verifies equation (4.8). $\square$

Exercise 4.6: Bloch sphere interpretation of rotations

$(*)$ One reason why the $R_{\hat{\mathbf{n}}}(\theta)$ operators are referred to as rotation operators is the following fact, which you are to prove. Suppose a single qubit has a state represented by the Bloch vector $\boldsymbol{\lambda}$. Then, the effects of the rotation $R_{\hat{\mathbf{n}}}(\theta)$ on the state is to rotate it by an angle θ about the $\hat{\mathbf{n}}$ axis of the Bloch sphere. This fact explains the rather mysterious looking factors of two in the definition of the rotation matrices.

Solution

Concepts Involved: Linear Algebra, Quantum Gates.

Let $\boldsymbol{\lambda}$ be an arbitrary Bloch vector. WLOG, we can express $\boldsymbol{\lambda}$ in a coordinate system such that $\hat{\mathbf{n}}$ is aligned with the $\hat{\mathbf{z}}$ axis, so it suffices to consider how the state behaves under application $R_z(\theta)$. Let $\boldsymbol{\lambda} = (\lambda_x, \lambda_y, \lambda_z)$ be the vector expressed in this coordinate system. By Exercise 2.72, the density operator corresponding to this Bloch vector is given by:

$$\rho = \frac{I + \boldsymbol{\lambda} \cdot \boldsymbol{\sigma}}{2}$$

We now observe how ρ transforms under conjugation by $R_z(\theta)$:

$$\begin{aligned} R_z(\theta)\rho R_z(\theta)^\dagger &= R_z(\theta)\rho R_z(-\theta) \\ &= R_z(\theta)\left(\frac{I + \lambda_x X + \lambda_y Y + \lambda_z Z}{2}\right)R_z(-\theta) \end{aligned}$$

Using that $XZ = -ZX$ from Exercise 2.41, we make the observation that:

$$\begin{aligned} R_z(\theta)X &= \left(\cos\left(\frac{\theta}{2}\right)I - i\sin\left(\frac{\theta}{2}\right)Z\right)X \\ &= X\left(\cos\left(\frac{\theta}{2}\right)I + i\sin\left(\frac{\theta}{2}\right)Z\right) \\ &= X\left(\cos\left(\frac{-\theta}{2}\right)I - i\sin\left(\frac{-\theta}{2}\right)Z\right) \\ &= XR_z(-\theta) \end{aligned}$$

Similarly, we find that $R_z(\theta)Y = R_z(-\theta)Y$ (same anticommutation) and that $R_z(\theta)Z = ZR_z(\theta)$ (all terms commute). With this, the expression for $R_z(\theta)\rho R_z(\theta)^\dagger$ simplifies to:

$$\begin{aligned} R_z(\theta)\rho R_z(\theta)^\dagger &= R_z(\theta)\left(\frac{I + \lambda_x X + \lambda_y Y + \lambda_z Z}{2}\right)R_z(-\theta) \\ &= \left(\frac{IR_z(\theta) + \lambda_x XR_z(-\theta) + \lambda_y YR_z(-\theta) + \lambda_z ZR_z(\theta)}{2}\right)R_z(-\theta) \\ &= \frac{I + \lambda_x XR_z(-2\theta) + \lambda_y YR_z(-2\theta) + \lambda_z Z}{2} \end{aligned}$$

Calculating each of the terms in the above expression, we have:

$$\begin{aligned} XR_z(-2\theta) &= X\left(\cos\left(\frac{-2\theta}{2}\right) - i\sin\left(\frac{-2\theta}{2}\right)Z\right) \\ &= X\big(\cos(\theta) + i\sin(\theta)Z\big) \\ &= \cos(\theta)X + i\sin(\theta)XZ \\ &= \cos(\theta)X + i\sin(\theta)(-iY) \\ &= \cos(\theta)X + \sin(\theta)Y \end{aligned}$$

$$
\begin{aligned}
YR_z(-2\theta) &= Y\left(\cos(\theta) + i\sin(\theta)Z\right) \\
&= \cos(\theta)Y + i\sin(\theta)YZ \\
&= \cos(\theta)Y + i\sin(\theta)(iX) \\
&= \cos(\theta)Y - \sin(\theta)X.
\end{aligned}
$$

Plugging these back into the expression for $R_z(\theta)\rho R_z(\theta)^\dagger$ and collecting like terms, we have:

$$
R_z(\theta)\rho R_z(\theta)^\dagger = \frac{I + (\lambda_x\cos(\theta) - \lambda_y\sin(\theta))X + (\lambda_x\sin(\theta) + \lambda_y\cos(\theta))Y + \lambda_z Z}{2}.
$$

From this expression, we can read off the new Bloch vector $\boldsymbol{\lambda}'$ after conjugation by $R_z(\theta)$ to be:

$$
\boldsymbol{\lambda}' = (\lambda_x\cos(\theta) - \lambda_y\sin(\theta), \lambda_x\sin(\theta) + \lambda_y\cos(\theta), \lambda_z).
$$

Alternatively, suppose we apply the 3-dimensional rotation matrix $A_z(\theta)$ to the original bloch vector $\boldsymbol{\lambda}$. We have that:

$$
A_z(\theta)\boldsymbol{\lambda} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda_x \\ \lambda_y \\ \lambda_z \end{bmatrix} = \begin{bmatrix} \lambda_x\cos\theta - \lambda_y\sin\theta \\ \lambda_x\sin\theta + \lambda_y\cos\theta \\ \lambda_z \end{bmatrix}.
$$

We see that we end up with the same resulting vector $\boldsymbol{\lambda}'$. We conclude that the conjugation of ρ under $R_z(\theta)$ has the equivalent effect to rotating the Bloch vector by θ about the $\hat{\mathbf{z}}$-axis, and hence the effect of $R_{\hat{\mathbf{n}}}(\theta)$ on a one qubit state is to rotate it by an angle θ about $\hat{\mathbf{n}}$. $\square$

Exercise 4.7

Show that $XYX = -Y$ and use this to prove that $XR_y(\theta)X = R_y(-\theta)$.

Solution

Concepts Involved: Linear Algebra, Quantum Gates.

For the first claim, we use that $XY = -YX$ and $X^2 = I$ (Exercise 2.41) to obtain that:

$$
XYX = -YXX = -YI = -Y.
$$

Using this, we have that:

$$
\begin{aligned}
XR_y(\theta)X &= X\left(\cos\left(\frac{\theta}{2}\right)I - i\sin\left(\frac{\theta}{2}\right)Y\right)X \\
&= \cos\left(\frac{\theta}{2}\right)XIX - i\sin\left(\frac{\theta}{2}\right)XYX \\
&= \cos\left(\frac{\theta}{2}\right)I + i\sin\left(\frac{\theta}{2}\right)Y \\
&= \cos\left(-\frac{\theta}{2}\right)I - i\sin\left(-\frac{\theta}{2}\right)Y \\
&= R_y(-\theta).
\end{aligned}
$$

□

Exercise 4.8

An arbitrary single qubit unitary operator can be written in the form

$$U = \exp(i\alpha)R_{\hat{\mathbf{n}}}(\theta)$$

for some real numbers α and θ, and a real three-dimensional unit vector $\hat{\mathbf{n}}$.

1. Prove this fact.
2. Find values for α, θ, and $\hat{\mathbf{n}}$ giving the Hadamard gate H.
3. Find values for α, θ, and $\hat{\mathbf{n}}$ giving the phase gate

$$S = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$$

Solution

Concepts Involved: Linear Algebra, Unitary Operators, Quantum Gates

---

1. By definition, for any unitary operator U we have that $U^\dagger U = I$, so for any state vector $\langle\psi|\psi\rangle = \langle\psi|U^\dagger U|\psi\rangle$. Therefore, all unitary Us are norm-preserving, and hence for a single qubit correspond to some reflection/rotation in 3-dimensional space (up to a global phase factor). Hence, we can write $U = \exp(i\alpha)R_{\hat{\mathbf{n}}}(\theta)$ for some $\hat{\mathbf{n}}$ (rotation axis), θ (rotation angle) and α (global phase).
2. Using the fact that $H = \frac{X+Z}{\sqrt{2}}$, and that modulo a factor of i that X/Z correspond to rotations

$R_x(\pi)$ and $R_z(\pi)$, we find that:

$$H = \frac{iR_x(\pi) + iR_z(\pi)}{\sqrt{2}} = i\left(\frac{2\cos\left(\frac{\pi}{2}\right)I - i\sin\left(\frac{\pi}{2}\right)X - i\sin\left(\frac{\pi}{2}\right)Z}{\sqrt{2}}\right)$$
$$= i\left(\cos\left(\frac{\pi}{2}\right)I - i\sin\left(\frac{\pi}{2}\right)\left(\frac{1}{\sqrt{2}}X + 0Y + \frac{1}{\sqrt{2}}Z\right)\right)$$
$$= e^{i\pi/2}\left(\cos\left(\frac{\pi}{2}\right)I - i\sin\left(\frac{\pi}{2}\right)\left(\frac{1}{\sqrt{2}}X + 0Y + \frac{1}{\sqrt{2}}Z\right)\right)$$

Note that in the second last equality we use that $\cos\left(\frac{\pi}{2}\right) = 0$ and hence $\frac{2}{\sqrt{2}}\cos\left(\frac{\pi}{2}\right) = \cos\left(\frac{\pi}{2}\right)$. From the last expression, we can read off using the definition of $R_{\hat{\mathbf{n}}}(\theta)$ that $\hat{\mathbf{n}} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$, $\theta = \pi$, and $\alpha = \frac{\pi}{2}$.

3. We observe that:

$$R_z\left(\frac{\pi}{2}\right) = \begin{bmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{bmatrix} = e^{-i\pi/4}\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$$

Hence:

$$S = e^{i\pi/4}R_z\left(\frac{\pi}{2}\right)$$

from which we obtain that $\hat{\mathbf{n}} = \hat{\mathbf{z}} = (0, 0, 1)$, $\theta = \frac{\pi}{2}$, and $\alpha = \frac{\pi}{4}$.

□

Remark: For part (2), one just can use the definition

$$R_{\hat{n}}(\theta) \equiv \exp\left(-i\theta\hat{n}\cdot\vec{\sigma}/2\right) = \cos\left(\frac{\theta}{2}\right)I - i\sin\left(\frac{\theta}{2}\right)\left(n_xX + n_yY + n_zZ\right),$$

and the fact $H = (X+Z)/\sqrt{2}$, to arrive at $\cos\left(\frac{\theta}{2}\right) = 0$, $n_x = n_z = \frac{1}{\sqrt{2}}, n_y = 0$.

Exercise 4.9

Explain why any single qubit unitary operator may be written in the form (4.12).

Solution

Concepts Involved: Linear Algebra, Unitary Operators, Quantum Gates.
Recall that (4.12) states that we can write any single qubit unitary U as:

$$U = \begin{bmatrix} e^{i(\alpha-\beta/2-\delta/2)}\cos\frac{\gamma}{2} & -e^{i(\alpha-\beta/2+\delta/2)}\sin\frac{\gamma}{2} \\ e^{i(\alpha+\beta/2-\delta/2)}\sin\frac{\gamma}{2} & e^{i(\alpha+\beta/2+\delta/2)}\cos\frac{\gamma}{2} \end{bmatrix}$$

where $\alpha, \beta, \gamma, \delta \in \mathbb{R}$.

Let U be a single qubit unitary operator. We then have that $U^\dagger U = I$, so identifying:

$$U = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}$$

we obtain that:

$$\begin{bmatrix} |a|^2 + |c|^2 & a^*b + c^*d \\ ab^* + cd^* & |b|^2 + |d|^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

From the diagonal entries we obtain that $|\mathbf{v}_1| = |\mathbf{v}_2| = 1$ and from the off diagonal entries we obtain that $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$ and hence the columns of U are orthonormal. From the fact that $|v|_1$ is normalized, we can parameterize the magnitude of the entries with $\gamma \in \mathbb{R}$ such that:

$$|a| = \cos\frac{\gamma}{2}, \quad |c| = \sin\frac{\gamma}{2}.$$

From the orthogonality, we further obtain that $b = -c^*$ and $d = a^*$, from which we have that $|b| = |c|$ and $|d| = |a|$. Furthermore, (also from the orthogonality) we can parameterize $\arg(a) = -\frac{\beta}{2} - \frac{\delta}{2}$ and $\arg(b) = \frac{\beta}{2} - \frac{\delta}{2}$ For $\beta, \delta \in \mathbb{R}$. Finally, multiplying U by a complex phase $e^{i\alpha}$ for $\alpha \in \mathbb{R}$ preserves the unitarity of U and the orthonormality of the colums. Combining these facts gives the form of (4.12) as desired. □

Exercise 4.10: $X - Y$ decomposition of rotations

Give a decomposition analogous to Theorem 4.1 but using R_x instead of R_z.

Exercise 4.11

Suppose $\hat{\mathbf{m}}$ and $\hat{\mathbf{n}}$ are non-parallel real unit vectors in three dimensions. Use Theorem 4.1 to show that an arbitrary single qubit unitary U may be written

$$U = e^{i\alpha} R_{\hat{\mathbf{n}}}(\beta) R_{\hat{\mathbf{m}}}(\gamma) R_{\hat{\mathbf{n}}}(\delta)$$

Exercise 4.12

Give A, B, C, and α for the Hadamard gate.

Solution

Concepts Involved: Linear Algebra, Decomposition of Rotations.
Recall that any single qubit unitary U can be written as $U = e^{i\alpha} AXBXC$ where $ABC = I$ and $\alpha \in \mathbb{R}$.

First, observe that we can write:

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = i \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = e^{i\pi/2} R_z(\pi) R_y(-\pi/2) R_z(0)$$

so defining A, B, C according to the proof of Corollary 4.2, we have:

$$\begin{aligned}A &= R_z(\pi)R_y(-\pi/4)\\ B &= R_y(\pi/4)R_z(-\pi/2)\\ C &= R_z(-\pi/2)\end{aligned}$$

and $\alpha = \frac{\pi}{2}$. □

Exercise 4.13: Circuit identities

It is useful to be able to simplify circuits by inspection, using well-known identities. Prove the following three identities:

$$HXH = Z; \quad HYH = -Y; \quad HZH = X.$$

Solution

Concepts Involved: Linear Algebra, Quantum Gates.

By computation, we find:

$$\begin{aligned}HXH &= \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}2 & 0\\ 0 & -2\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} = Z\\ HYH &= \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}0 & -2i\\ 2i & 0\end{bmatrix} = -\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix} = -Y\\ HZH &= \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}0 & 2\\ 2 & 0\end{bmatrix} = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} = X\end{aligned}$$

□

Remark: Notice once we have proved $HXH = Z$, we can directly say $HZH = H(HXH)H = X$ as $H^2 = I$. If one wants to prove everything algebraically, the following calculation suffices.

$$HXH := \frac{1}{2}(X+Z)X(X+Z) = \frac{1}{2}(I+ZX)(X+Z) = \frac{1}{2}(X+Z+Z+XZX) = Z$$
$$HYH := \frac{1}{2}(X+Z)Y(X+Z) = \frac{1}{2}(XY+ZY)(X+Z) = \frac{1}{2}(XYX+ZXY+ZYX+ZYZ) = -Y$$

Exercise 4.14

Use the previous exercise to show that $HTH = R_x(\pi/4)$, up to a global phase.

Solution

Concepts Involved: Linear Algebra, Quantum Gates.

From Exercise 4.3, we know that $T = R_z(\pi/4)$ up to a global phase $e^{-i\pi/8}$. We hence have that:

$$\begin{aligned}
HTH &= e^{-i\pi/8} H R_z(\pi/4) H \\
&= e^{-i\pi/8} H \left(\cos\left(\frac{\pi}{8}\right) I - i\sin\left(\frac{\pi}{8}\right) Z\right) H \\
&= e^{-i\pi/8} \left(\cos\left(\frac{\pi}{8}\right) I - i\sin\left(\frac{\pi}{8}\right) X\right) \\
&= e^{-i\pi/8} R_x(\pi/4)
\end{aligned}$$

where in the second last equality we use the previous exercise, as well as the fact that $HIH = H^2 = I$ from Exercise 2.52. □

Exercise 4.15: Composition of single qubit operations

The Bloch representation gives a nice way to visualize the effect of composing two rotations.

(1) Prove that if a rotation through an angle β_1 about the axis $\hat{\mathbf{n}}_1$ is followed by a rotation through an angle β_2 about an axis $\hat{\mathbf{n}}_2$, then the overall rotation is through an angle β_{12} about an axis $\hat{\mathbf{n}}_{12}$ given by

$$\begin{aligned}
c_{12} &= c_1 c_2 - s_1 s_2 \hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2 \\
s_{12}\hat{\mathbf{n}}_{12} &= s_1 c_2 \hat{\mathbf{n}}_1 + c_1 s_2 \hat{\mathbf{n}}_2 - s_1 s_2 \hat{\mathbf{n}}_2 \times \hat{\mathbf{n}}_1,
\end{aligned}$$

where $c_i = \cos\big(\beta_i/2\big)$, $s_i = \sin\big(\beta_i/2\big)$, $c_{12} = \cos\big(\beta_{12}/2\big)$, and $s_{12} = \sin\big(\beta_{12}/2\big)$.

(2) Show that if $\beta_1 = \beta_2$ and $\hat{\mathbf{n}}_1 = \hat{\mathbf{z}}$ these equations simplify to

$$\begin{aligned}
c_{12} &= c^2 - s^2 \hat{\mathbf{z}} \cdot \hat{\mathbf{n}}_2 \\
s_{12}\hat{\mathbf{n}}_{12} &= sc(\hat{\mathbf{z}} + \hat{\mathbf{n}}_2) - s^2 \hat{\mathbf{n}}_2 \times \hat{\mathbf{z}}
\end{aligned}$$

Solution

Concepts Involved: Linear Algebra

(1) It suffices to show that $R_{\hat{\mathbf{n}}_2}(\beta_2) R_{\hat{\mathbf{n}}_1}(\beta_1)$ is equivalent to $R_{\hat{\mathbf{n}}_{12}}(\beta_{12})$.

$$\begin{aligned}
R_{\hat{\mathbf{n}}_2}(\beta_2) R_{\hat{\mathbf{n}}_1}(\beta_1) &= (c_2 I - i s_2 \hat{\mathbf{n}}_2 \cdot \boldsymbol{\sigma}) \cdot (c_1 I - i s_1 \hat{\mathbf{n}}_1 \cdot \hat{\sigma}) \\
&= c_2 c_1 I - i\left(c_1 s_2 \hat{\mathbf{n}}_2 \cdot \boldsymbol{\sigma} + c_2 s_1 \hat{\mathbf{n}}_1 \cdot \boldsymbol{\sigma}\right) - s_2 s_1 \underbrace{(\hat{\mathbf{n}}_2 \cdot \boldsymbol{\sigma}) \cdot (\hat{\mathbf{n}}_1 \cdot \boldsymbol{\sigma})}_{(\hat{\mathbf{n}}_2 \cdot \hat{\mathbf{n}}_1) I + i(\hat{\mathbf{n}}_2 \times \hat{\mathbf{n}}_1) \cdot \boldsymbol{\sigma}} \\
&= \left[c_2 c_1 - s_2 s_1 \left(\hat{\mathbf{n}}_2 \cdot \hat{\mathbf{n}}_1\right)\right] I - i\left[c_1 s_2 \hat{\mathbf{n}}_\mathbf{2} + c_2 s_1 \hat{\mathbf{n}}_\mathbf{1} + s_2 s_1 \left(\hat{\mathbf{n}}_2 \times \hat{\mathbf{n}}_1\right)\right] \cdot \boldsymbol{\sigma}
\end{aligned}$$

Identifying this operation to a single rotation $R_{\hat{\mathbf{n}}_{12}}(\beta_{12}) \equiv c_{12} I - i s_{12} \hat{\mathbf{n}}_{12} \cdot \boldsymbol{\sigma}$, we arrive at the required relations (up to a presumable typesetting error)

$$
\begin{aligned}
c_{12} &= c_2 c_1 - s_2 s_1 \left(\hat{\mathbf{n}}_2 \cdot \hat{\mathbf{n}}_1\right) \\
s_{12}\hat{\mathbf{n}}_{12} &= c_1 s_2 \hat{\mathbf{n}}_{\mathbf{2}} + c_2 s_1 \hat{\mathbf{n}}_{\mathbf{1}} + s_2 s_1 \left(\hat{\mathbf{n}}_2 \times \hat{\mathbf{n}}_1\right)
\end{aligned}
$$

(2) Setting $\beta_1 = \beta_2$ and $\hat{\mathbf{n}}_1 = \hat{\mathbf{z}}$ in the formulas proven above combined with the fact that $c = c_1 = \cos(\beta_1/2) = \cos(\beta_2/2) = c_2$ (and similiarly $s = s_1 = s_2$), we have:

$$
\begin{aligned}
c_{12} &= c^2 - s^2 \hat{\mathbf{z}} \cdot \hat{\mathbf{n}}_2 \\
s_{12}\hat{\mathbf{n}}_{12} &= sc\hat{\mathbf{z}} + cs\hat{\mathbf{n}}_2 - s^2 \hat{\mathbf{n}}_2 \times \hat{\mathbf{z}} = sc(\hat{\mathbf{z}} + \hat{\mathbf{n}}_2) - s^2 \hat{\mathbf{n}}_2 \times \hat{\mathbf{z}}.
\end{aligned}
$$

□

Remark: For the sake of completeness, we provide a proof of the identity used in part 1 of the solution. First note the familiar Pauli matrix relation $\sigma_i \sigma_j = \delta_{ij} I + i\epsilon_{ijk}\sigma_k$ (Exercise 2.43). Now massaging this equation gives

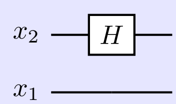
$$
\begin{aligned}
a_i \sigma_i b_j \sigma_j &= a_i b_j \delta_{ij} + i \left(a_i b_j \epsilon_{ijk}\right) \sigma_k \\
&= (\mathbf{a} \cdot \mathbf{b})\, I + i\, (\mathbf{a} \times \mathbf{b})_k\, \sigma_k,
\end{aligned}
$$

where we have used standard Einstein index notation. Thus in matrix form, we have

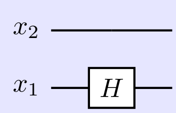
$$
(\mathbf{a} \cdot \boldsymbol{\sigma}) \cdot (\mathbf{b} \cdot \boldsymbol{\sigma}) = (\mathbf{a} \cdot \mathbf{b})\, I \, + \, i\, (\mathbf{a} \times \mathbf{b}) \cdot \boldsymbol{\sigma}\,.
$$

Exercise 4.16

What is the 4×4 unitary matrix for the circuit



in the computational basis? What is the unitary matrix for the circuit



Solution

Concepts Involved: Linear Algebra, Quantum Gates, Tensor Products.

The unitary matrix for the first circuit is given by:

$$
I_1 \otimes H_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}.
$$

The unitary matrix for the second circuit is given by:

$$H_1 \otimes I_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

□

Exercise 4.17: Building a `CNOT` from controlled-Z gates

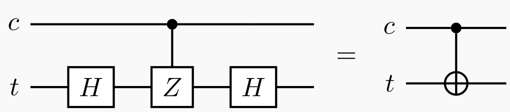
Construct a `CNOT` gate from one controlled-Z gate, that is, the gate whose action in the computational basis is specified by the unitary matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Solution

Concepts Involved: Linear Algebra, Quantum Gates, Controlled Operations.

We showed in Exercise 4.13 that $HZH = X$. Hence, to obtain a CNOT gate from a single controlled Z gate, we can conjugate the target qubit with Hadamard gates:



We can verify this via matrix multiplication, using the result from the previous exercise:

$$\begin{aligned}
(I_1 \otimes H_2)(CZ_{1,2})(I_1 \otimes H_2) &= \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \\
&= \frac{1}{2} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix} \\
&= CX_{1,2}
\end{aligned}$$

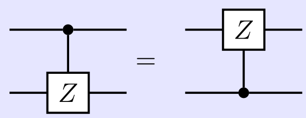
□

Remark:

$$\begin{aligned} CX_{1,2} &:= |0\rangle\langle 0| \otimes I + |0\rangle\langle 0| \otimes X \\ &= |0\rangle\langle 0| \otimes HH + |0\rangle\langle 0| \otimes HZH \\ &=: (I \otimes H)(CZ_{1,2})(I \otimes H). \end{aligned}$$

Exercise 4.18

Show that



Solution

Concepts Involved: Linear Algebra, Quantum Gates, Controlled Operations.

It suffices to verify that the two gates have the same effect on the 2-qubit computational basis states (as it will then follow by linearity that they will have the same effect on any such superposition of the basis states). Checking the 8 necessary cases, we then have that:

$$\begin{aligned} &CZ_{1,2}(|0\rangle_1 \otimes |0\rangle_2) = |0\rangle_1 \otimes |0\rangle_2 \\ &CZ_{2,1}(|0\rangle_1 \otimes |0\rangle_2) = |0\rangle_1 \otimes |0\rangle_2 \\ &CZ_{1,2}(|1\rangle_1 \otimes |0\rangle_2) = |1\rangle_1 \otimes Z|0\rangle_2 = |1\rangle_1 \otimes |0\rangle_2 \\ &CZ_{2,1}(|1\rangle_1 \otimes |0\rangle_2) = |1\rangle_1 \otimes |0\rangle_2 \\ &CZ_{1,2}(|0\rangle_1 \otimes |1\rangle_2) = |0\rangle_1 \otimes |1\rangle_2 \\ &CZ_{2,1}(|0\rangle_1 \otimes |1\rangle_2) = Z|0\rangle_1 \otimes |1\rangle_2 = |0\rangle_1 \otimes |1\rangle_2 \\ &CZ_{1,2}(|1\rangle_1 \otimes |1\rangle_2) = |1\rangle_1 \otimes Z|1\rangle_2 = |1\rangle_1 \otimes -|1\rangle_2 = -(|1\rangle_1 \otimes |1\rangle_1) \\ &CZ_{2,1}(|1\rangle_1 \otimes |1\rangle_2) = Z|1\rangle_1 \otimes |1\rangle_2 = -|1\rangle_1 \otimes |1\rangle_2 = -(|1\rangle_1 \otimes |1\rangle_1) \end{aligned}$$

from which we observe equality for each. The claim follows. □

Remark: More compactly, we have $CZ_{1,2}|b_1b_2\rangle = |b_1\rangle \otimes Z^{b_1}|b_2\rangle = (-1)^{b_1 \cdot b_2}|b_1b_2\rangle$ for computational basis states $b_1, b_2 \in \{0, 1\}$.
Using this form we can write

$$\begin{aligned} CZ_{1,2}|b_1b_2\rangle &= (-1)^{b_1 \cdot b_2}|b_1b_2\rangle \\ &= (-1)^{b_2 \cdot b_1}|b_1b_2\rangle \\ &= Z^{b_2}|b_1\rangle \otimes |b_2\rangle \\ &=: CZ_{2,1}|b_1b_2\rangle. \end{aligned}$$

Exercise 4.19: `CNOT` action on unitary matrices

The `CNOT` gate is a simple permutation whose action on a density matrix ρ is to rearrange the elements in the matrix. Write out this action explicitly in the computational basis.

Solution

Concepts Involved: Linear Algebra, Quantum Gates, Controlled Operations, Density Operators

Let ρ be an arbitrary density matrix corresponding to a 2-qubit state. In the computational basis, we can write ρ as:

$$\rho \cong \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}.$$

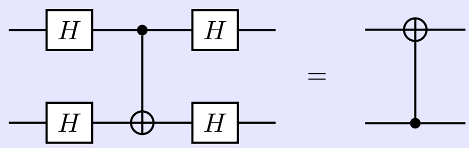
Studying the action of the `CNOT` gate on this density matrix, we calculate:

$$\begin{aligned} CX_{1,2}\,\rho\, CX_{1,2} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{41} & a_{42} & a_{43} & a_{34} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & a_{14} & a_{13} \\ a_{21} & a_{22} & a_{24} & a_{23} \\ a_{41} & a_{42} & a_{44} & a_{33} \\ a_{31} & a_{32} & a_{34} & a_{33} \end{bmatrix} \end{aligned}$$

$\square$

Exercise 4.20: `CNOT` **basis transformations**

Unlike ideal classical gates, ideal quantum gates do not have (as electrical engineers say) 'high-impedance' inputs. In fact, the role of 'control' and 'target' are arbitrary – they depend on what basis you think of a device as operating in. We have described how the `CNOT` behaves with respect to the computational basis, and in this description the state of the control qubit is not changed. However, if we work in a different basis then the control qubit *does* change: we will show that its phase is flipped depending on the state of the 'target' qubit! Show that



Introducing basis states $|\pm\rangle \equiv (|0\rangle \pm |1\rangle)/\sqrt{2}$, use this circuit identity to show that the effect of a `CNOT` with the first qubit as control and the second qubit as target is as follows:

$$
\begin{aligned}
|+\rangle|+\rangle &\mapsto |+\rangle|+\rangle \\
|-\rangle|+\rangle &\mapsto |-\rangle|+\rangle \\
|+\rangle|-\rangle &\mapsto |-\rangle|-\rangle \\
|-\rangle|-\rangle &\mapsto |+\rangle|-\rangle.
\end{aligned}
$$

Thus, with respect to this new basis, the state of the target qubit is not changed, while the state of the control qubit is flipped if the target starts as $|-\rangle$, otherwise it is left alone. That is, in this basis, the target and control have essentially interchanged roles!

Solution

Concepts Involved: Linear Algebra, Quantum Gates, Controlled Operations.

First, we have that:

$$H_1 \otimes H_2 = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

Now conjugating $\mathsf{CNOT}_{1,2}$ under $H_1 \otimes H_2$, we have:

$$
\begin{aligned}
(H_1 \otimes H_2)CX_{1,2}(H_1 \otimes H_2) &= \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \\
&= \frac{1}{4}\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 4 & 0 \\ 0 & 4 & 0 & 0 \end{bmatrix} \\
&= CX_{2,1}
\end{aligned}
$$

which proves the circuit identity. We know already that:

$$
\begin{aligned}
CX_{2,1}|0\rangle|0\rangle &= |0\rangle|0\rangle \\
CX_{2,1}|1\rangle|0\rangle &= |1\rangle|0\rangle \\
CX_{2,1}|0\rangle|1\rangle &= |1\rangle|1\rangle \\
CX_{2,1}|1\rangle|1\rangle &= |0\rangle|1\rangle
\end{aligned}
$$

so using the proven circuit identity and the fact that $H|0\rangle = |+\rangle, H|1\rangle = |-\rangle$, we obtain the map:

$$
\begin{aligned}
|+\rangle|+\rangle &\mapsto |+\rangle|+\rangle \\
|-\rangle|+\rangle &\mapsto |-\rangle|+\rangle \\
|+\rangle|-\rangle &\mapsto |-\rangle|-\rangle \\
|-\rangle|-\rangle &\mapsto |+\rangle|-\rangle
\end{aligned}
$$

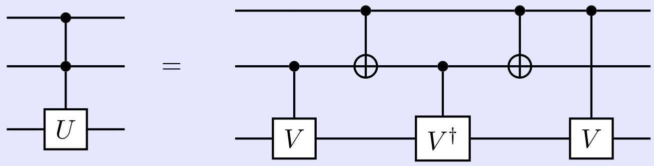
which is exactly what we wanted to prove. $\square$

Remark: Algebraically,

$$
\begin{aligned}
(H \otimes H)CX_{1,2}(H \otimes H) &= (H \otimes H)(I \otimes H)(CZ_{1,2})(I \otimes H)(H \otimes H) \\
&= (H \otimes I)(CZ_{1,2})(H \otimes I) \\
&= (H \otimes I)(CZ_{2,1})(H \otimes I) \\
&= CX_{2,1}.
\end{aligned}
$$

Exercise 4.21

Verify that Figure 4.8 implements the $C^2(U)$ operation.



Exercise 4.22

Prove that a $C^2(U)$ gate (for any single qubit unitary U) can be constructed using at most eight one-qubit gates, and six controlled-`NOT`s.

Exercise 4.23

Construct a $C^1(U)$ gate for $U = R_x(\theta)$ and $U = R_y(\theta)$, using only `CNOT` and single qubit gates. Can you reduce the number of single qubit gates needed in the construction from three to two?

Exercise 4.24

Verify that Figure 4.9 implements the Toffoli gate.

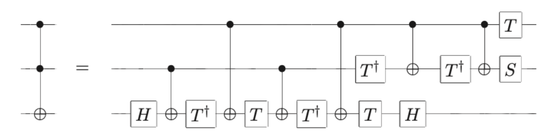


Figure 4.9. Implementation of the Toffoli gate using Hadamard, phase, controlled-NOT and $\pi/8$ gates.

Exercise 4.25: Fredkin gate construction

Recall that the Fredkin (controlled-swap) gate performs the transform

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(1) Give a quantum circuit which uses three Toffoli gates to construct the Fredkin gate (*Hint:* think of the swap gate construction – you can control each gate, one at a time).

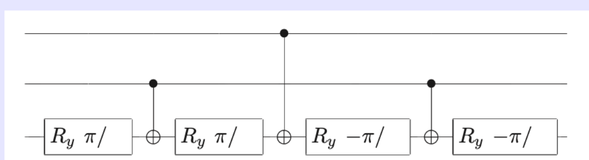
(2) Show that the first and last Toffoli gates can be replaced by `CNOT` gates.

(3) Now replace the middle Toffoli gate with the circuit in Figure 4.8 to obtain a Fredkin gate construction using only six two-qubit gates.

(4) Can you come up with an even simpler construction, with only five two-qubit gates?

Exercise 4.26

Show that the circuit:



differs by a Toffoli gate only by relative phases. That is, the circuit that takes $|c_1, c_2, t\rangle$ to $e^{i\theta(c_1,c_2,t)}\,|c_1, c_2, t \oplus c_1 \cdot c_2\rangle$, where $e^{i\theta(c_1,c_2,t)}$ is some relative phase factor. Such gates can be sometimes be useful in experimental implementations, where it may be much easier to implement a gate that is the same as the Toffoli gate up to relative phases than it is to do the Toffoli directly.

Exercise 4.27

Using just CNOTs and Toffoli gates, construct a quantum circuit to perform the transformation

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

This kind of partial cyclic permutation operation will be useful later, in Chapter 7.

Exercise 4.28

For $U = V^2$ with V unitary, construct a $C^5(U)$ gate analogous to that in Figure 4.10, but using no work qubits. You may use controlled-V and controlled-$V^\dagger$ gates.

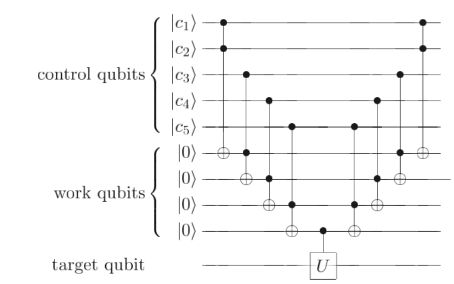


Figure 4.10. Network implementing the $C^n(U)$ operation, for the case $n = 5$.

Exercise 4.29

Find a circuit containing $O(n^2)$ Toffoli, CNOT and single qubit gates which implements a $C^n(X)$ gate (for $n > 3$), using no work qubits.

Exercise 4.30

Suppose U is a single qubit unitary operation. Find a circuit containing $O(n^2)$ Toffoli, CNOT and single qubit gates which implements a $C^n(U)$ gate (for $n > 3$), using no work qubits.

Exercise 4.31: More circuit identities

Let subscripts denote which qubit an operator acts on, and ket C be a `CNOT` with qubit 1 the control qubit and qubit 2 the target qubit. Prove the following identities:

$$\begin{aligned}
CX_1C &= X_1X_2 \\
CY_1C &= Y_1X_2 \\
CZ_1C &= Z_1 \\
CX_2C &= X_2 \\
CY_2C &= Z_1Y_2 \\
CZ_2C &= Z_1Z_2 \\
R_{z,1}(\theta)C &= CR_{z,1}(\theta) \\
R_{x,2}(\theta)C &= CR_{x,2}(\theta).
\end{aligned}$$

Exercise 4.32

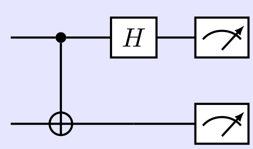
Let ρ be the density matric describing a two qubit system. Suppose we perform a projective measurement in the computational basis of the second qubit. Let $P_0 = |0\rangle\langle 0|$ and $P_1 = |1\rangle\langle 1|$ be the projectors onto the $|0\rangle$ and the $|1\rangle$ states of the second qubit, respectively. Let ρ' be the density matrix which would be assigned to the system after the measurement by an observer who did not learn the measurement result. Show that

$$\rho' = P_0\rho P_0 + P_1\rho P_1$$

Also show that the reduced density matrix for the first qubit is not affected by the measurement, that is $\mathrm{tr}_2(\rho) = \mathrm{tr}_2(\rho')$.

Exercise 4.33: Measurement in the Bell basis

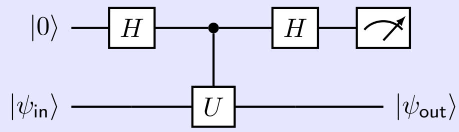
The measurement model we have specified for the quantum circuit model is that measurements are performed only in the computational basis. However, often we want to perform a measurement in some other basis, defined by a complete set of orthonormal states. To perform this measurement, simply unitarily transform from the basis we wish to perform the measurement in to the computational basis, then measure. For example, show that the circuit



performs a measurement in the basis of the Bell states. More precisely, show that this circuit results in a measurement being performed with corresponding POVM elements the four projectors onto the Bell states. What are the corresponding measurement operators?

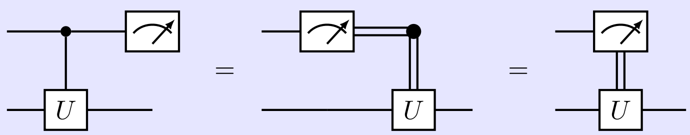
Exercise 4.34: Measuring an operator

Suppose we have a single qubit operator U with eigenvalues ± 1, so that U is both Hermitian and unitary, so it can be regarded as both an observable and a quantum gate. Suppose we wish to measure the observable U. That is, we desire to obtain a measurement result indicating one of the two eigenvalues, and leaving a post-measurement state which is the corresponding eigenvector. How can this be implemented by a quantum circuit? Show that the following circuit implements a measurement of U:



Exercise 4.35: Measurement commutes with controls

A consequence of the principle of deferred measurement is that measurements commute with quantum gates when the qubit being measured is a control qubit, that is:



(Recall that the double lines represent classical bits in this diagram.) Prove the first equality. The rightmost circuit is simply a convenient notation to depict the use of a measurement result to classically control a quantum gate.

Exercise 4.36

Construct a quantum circuit to add two two-bit numbers x and y modulo 4. That is, the circuit should perform the transformation $|x, y\rangle \mapsto |x, x + y \bmod 4\rangle$.

Exercise 4.37

Provide a decomposition of the transform

$$\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & 1 \\ 1 & -i & -1 & i \end{bmatrix}$$

into a product of two-level unitaries. This is a special case of the quantum Fourier transform, which we study in more detail in the next chapter.

Exercise 4.38

Prove that there exist a $d \times d$ unitary matrix U which cannot be decomposed as a product of fewer than $d - 1$ two-level unitary matrices.

Exercise 4.39

Find a quantum circuit using single qubit operations and CNOTs to implement the transformation

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 & 0 & 0 & c \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & b & 0 & 0 & 0 & 0 & d \end{bmatrix}$$

where $\tilde{U} = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$ is an arbitrary 2×2 unitary matrix.

Exercise 4.40

For arbitrary α and β show that

$$E(R_{\hat{\mathbf{n}}}(\alpha), R_{\hat{\mathbf{n}}}(\theta)^n) < \frac{\epsilon}{3}$$

and use this to justify (4.76).

Exercise 4.41

This and the next two exercises develop a construction showing that the Hadamard, phase, controlled-NOT and Toffoli gates are universal. Show that the circuit in Figure 4.17 applies the operation $R_z(\theta)$ to the third (target) qubit if the measurement outcomes are both 0, where $\cos\theta = 3/5$, and otherwise applies Z to the target qubit. Show that the probability of both measurement outcomes being 0 is $5/8$, and explain how repeated use of this circuit and $Z = S^2$ gates may be used to apply a $R_z(\theta)$ gate with probability approaching 1.

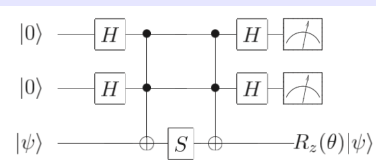


Figure 4.17. Provided both measurement outcomes are 0 this circuit applies $R_z(\theta)$ to the target, where $\cos\theta = 3/5$. If some other measurement outcome occurs then the circuit applies Z to the target.

Exercise 4.42: Irrationality of θ

Suppose $\cos\theta = 3/5$. We give a proof by contradiction that θ is an irrational multiple of 2π.

(1) Using the fact that $e^{i\theta} = (3+4i)/5$, show that if θ is rational, then there must exist a positive integer m such that $(3+4i)^m = 5^m$.

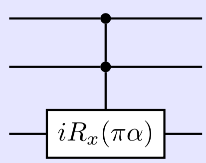
(2) Show that $(3+4i)^m = 3+4i \pmod 5$ for all $m > 0$, and conclude that no m such that $(3+4i)^m = 5^m$ can exist.

Exercise 4.43

Use the results of the previous two exercises to show that the Hadamard, phase, controlled-`NOT` and Toffoli gates are universal for quantum computation.

Exercise 4.44

Show that the three qubit gate G defined by the circuit:



is universal for quantum computation whenever α is irrational.

Exercise 4.45

Suppose U is a unitary transform implemented by an n qubit quantum circuit constructed from H, S, `CNOT` and Toffoli gates. Show that U is of the form $2^{-k/2}M$ for some integer k, where M is a $2^n \times 2^n$ matrix with only complex integer entries. Repeat this exercise with the TOffoli gate replaced by the $\pi/8$ gate.

Exercise 4.46: Exponential complexity growth of quantum systems

Let ρ be a density matrix describing the state of n qubits. Show that describing ρ requires $4^n - 1$ independent real numbers.

Exercise 4.47

For $H = \sum_k^L H_k$, prove that $e^{-iHt} = e^{-iH_1 t} e^{-iH_2 t} \dots e^{-iH_L t}$ for all i if $[H_j, H_k] = 0$, for all j, k.

Exercise 4.48

Show that the restriction of H_k to at most c particles *implies* that in the sum (4.97), L is upper bounded by a polynomial in n.

Exercise 4.49: Baker–Campbell–Hausdorf formula

Prove that

$$e^{(A+B)\Delta t} = e^{A\Delta t} e^{B\Delta t} e^{-\frac{1}{2}[A,B]\Delta t^2} + O(\Delta t^3)$$

and also prove Equations (4.103) and (4.104).

Exercise 4.50

Let $H = \sum_k^L H_k$, and define

$$U_{\Delta t} = \left[e^{-iH_1\Delta t} e^{-iH_2\Delta t} \dots e^{-iH_L\Delta t}\right]\left[e^{-iH_L\Delta t} e^{-iH_{L-1}\Delta t} \dots e^{-iH_1\Delta t}\right]$$

(a) Prove that $U_{\Delta t} = e^{-2iH\Delta t} + O(\Delta t^3)$

(b) Use the results in Box 4.1 to prove that for a positive integer m,

$$E(U^m_{\Delta t}, e^{-2miH\Delta t}) \leq m\alpha\Delta t^3,$$

for some constant α.

Exercise 4.51

Construct a quantum circuit to simulate the Hamiltonian

$$H = X_1 \otimes Y_2 \otimes Z_3$$

performing the unitary transform $e^{-i\Delta t H}$ for any Δt.

Problem 4.1: Computable phase shifts

Let m and n be positive integers. Suppose $f : \{0, \dots, 2^m - 1\} \mapsto \{0, \dots, 2^n - 1\}$ is a classical function from m to n bits which may be computed reversibly using T Toffoli gates, as described in Section 3.2.5. That is, the function $(x, y) \mapsto (x, y \oplus f(x))$ may be implemented using T Toffoli gates. Give a quantum circuit using $2T + n$ (or fewer) one, two and three qubit gates to implement the unitary operation defined by

$$|x\rangle \mapsto \exp\left(\frac{-2i\pi f(x)}{2^n}\right)|x\rangle$$

Problem 4.2

Find a depth $O(\log n)$ construction for the $C^n(X)$ gate. (*Comment:* The depth of a circuit is the number of distinct timesteps at which gates are applied; the point of this problem is that it is possible to parallelize the $C^n(X)$ construction by applying many gates in parllel during the same timestep.)

Problem 4.3: Alternate universality construction

Suppose U is a unitary matrix on n qubits. Define $H \equiv i \ln(U)$. Show that

(1) H is Hermitian, with eigenvalues in the range 0 to 2π.

(2) H can be written

$$H = \sum_g h_g g,$$

where h_g are real numbers and the sum is over all n-fold tensor products g of the Pauli matrices $\{I, X, Y, Z\}$.

(3) Let $\Delta = 1/k$, for some positive integer k. Explain how the unitary operation $\exp\big(-ih_g g\Delta\big)$ may be implemented using $O(n)$ one and two qubit operations.

(4) Show that

$$\exp(-iH\Delta) = \prod_g \exp\big(-ih_g g\Delta\big) + O(4^n\Delta^2)$$

where the product is taken with respect to any fixed ordering of the n-fold tensor products of Pauli matrices, g.

(5) Show that

$$U = \left[\prod_g \exp\big(-ih_g h\Delta\big)\right]^k + O(4^n\Delta)$$

(6) Explain how to approximate U to within a distance $\epsilon > 0$ using $O(n16^n/\epsilon)$ one and two qubit unitary operations.

Problem 4.4: Minimal Toffoli construction (Research)

The following problems concern constructions of the Toffoli with some minimal number of other gates. ‘

(1) What is the smallest number of two qubit gates that can be used to implement the Toffoli gate?

(2) What is the smallest number of one qubit gates and `CNOT` gates that can be used to implement the Toffoli gate?

(3) What is the smallest number of one qubit gates and controlled-Z gates that can be used to implement the Toffoli gate?

Problem 4.5: (Research)

Construct a family of Hamiltonians, $\{H_n\}$, on n qubits, such that simulating H_n requires a number of operations super-polynomial in n. (*Comment:* This problem seems to be quite difficult.)

Problem 4.6: Universality with prior entanglement

Controlled-`NOT` gates and single qubit gates form a universal set of quantum logic gates. Show that an alternative universal set of resources is comprised of single qubit unitaries, the ability to perform measurements of pairs of qubits in the Bell basis, and the ability to prepare arbitrary four qubit entangled states.

A1 Notes on basic probability theory

Exercise A1.1

Prove Bayes' rule.

Solution

Concepts Involved: Probability, Conditional Probability.
Recall that conditional probabilities were defined as:

$$P(Y = y|X = x) = \frac{P(X = x, Y = y)}{P(X = x)}$$

and also recall that Bayes' rule is given by:

$$p(x|y) = p(y|x)\frac{p(x)}{p(y)}$$

By the definition of conditional probability:

$$p(y|x)\frac{p(x)}{p(y)} = \frac{p(X = x, Y = y)}{p(x)}\frac{p(x)}{p(y)} = \frac{P(X = x, Y = y)}{p(y)} = p(x|y)$$

$\square$

Exercise A1.2

Prove the law of total probability.

Solution

Concepts Involved: Probability, Conditional Probability.
Recall that the law of total probability is given by:

$$p(y) = \sum_x p(y|x)p(x)$$

Using the identity $p(Y = y) = \sum_x p(X = x, Y = y)$ and Bayes' rule, we have

$$p(y) = \sum_x p(x, y) = \sum_x p(y|x)p(x)$$

$\square$

Exercise A1.3

Prove that there exists a value of $x \geq \mathbf{E}(X)$ such that $p(x) > 0$.

Solution

Concepts Involved: Probability, Expectation.

Recall that the expectation of a random variable X is defined by:

$$\mathbf{E}(X) = \sum_x p(x)x$$

Let $\tilde{x} = \max\{x : x \text{ is a possible value of } X\}$. This maximum exists as we assume X can only take on a finite set of values. We therefore have that:

$$\mathbf{E}(X) = \sum_x p(x)x \leq \sum_x p(x)\tilde{x} = \tilde{x}\sum_x p(x) = \tilde{x}$$

Where in the last equality we use that the sum over all probabilities must be 1. $\square$

Exercise A1.4

Prove that $\mathbf{E}(X)$ is linear in X

Solution

Concepts Involved: Probability, Expectation.

Let $a, b \in \mathbb{R}$ and X, Y be random variables. We then have that:

$$\begin{aligned}
\mathbf{E}(aX + bY) &= \sum_x \sum_y p(x,y)(ax + by) \\
&= \sum_x \sum_y p(x,y)ax + \sum_x \sum_y p(x,y)by \\
&= a\sum_x \left(\sum_y p(x,y)\right) x + b\sum_y \left(\sum_x p(x,y)\right) y \\
&= a\sum_x p(x)x + b\sum_y p(y)y \\
&= a\mathbf{E}(X) + b\mathbf{E}(y)
\end{aligned}$$

which shows that expectation is linear. $\square$

Exercise A1.5

Prove that for independent random variables X and Y, $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$.

Solution

Concepts Involved: Probability, Expectation, Independent Random Variables.

Recall two random variables X, Y are independent if

$$p(X = x, Y = y) = p(X = x)p(Y = y)$$

We have that:

$$\mathbf{E}(XY) = \sum_x \sum_y xyp(x, y) = \sum_x \sum_y xyp(x)p(y) = \left(\sum_x p(x)x\right)\left(\sum_y p(y)y\right) = \mathbf{E}(x)\mathbf{E}(y)$$

□

Exercise A1.6

$(*)$ Prove Chebyshev's inequality.

Solution

Concepts Involved: Probability, Expectation, Variance.
Recall the definition of the variance and standard deviaiton of a random variable X:

$$\mathrm{Var}(X) = \mathbf{E}[(X - \mathbf{E}(X))^2], \quad \Delta(X) = \sqrt{\mathrm{Var}(X)}$$

Also, recall that Chebyshev's inequality reads:

$$p(\left|X - \mathbf{E}(X)\right| \geq \lambda\Delta(X)) \leq \frac{1}{\lambda^2}$$

where $\lambda > 0$.

We first establish Markov's inequality for the expectation value $\mathbf{E}(X)$. Let $a > 0$, and then we have that:

$$\mathbf{E}(X) = \sum_x xp(x) = \sum_{x \geq a} xp(x) + \sum_{x<a} xp(x) \geq \sum_{x \geq a} ap(x) + 0 = ap(X \geq a)$$

Therefore, we obtain that:

$$p(X \geq a) \leq \frac{\mathbf{E}(X)}{a}$$

for any random variable X and $a > 0$. Next, substitute X with $(X - \mathbf{E}(X))^2$ and let $a = \lambda^2 \mathrm{Var}(X)$ for $\lambda > 0$. Markov's inequality then states that:

$$p((X - \mathbf{E}(X))^2 \geq \lambda^2 \mathrm{Var}(X)) \leq \frac{\mathbf{E}(X - \mathbf{E}(X))^2}{\lambda^2 \mathrm{Var}(X)}$$

Since $\mathbf{E}(X - \mathbf{E}(X))^2 = \mathrm{Var}(X)$, we have that:

$$p((X - \mathbf{E}(X))^2 \geq \lambda^2 \mathrm{Var}(X)) \leq \frac{1}{\lambda^2}$$

If $\lambda > 0$, then $p((X - \mathbf{E}(X))^2 \geq \lambda^2 \operatorname{Var}(X)) = p(\left|X - \mathbf{E}(X)\right| \geq \lambda\Delta(X))$ by taking square roots, so we obtain:

$$p(\left|X - \mathbf{E}(X)\right| \geq \lambda\Delta(X)) \leq \frac{1}{\lambda^2}$$

as desired. $\square$

A2 Group theory

Exercise A2.1

Prove that for any element of g of a finite group, there always exists a positive integer r such that $g^r = e$. That is, every element of such a group has an order.

Solution

Concepts Involved: Group Axioms, Order.

Suppose G is a finite group, and $g \in G$. Then, there exists some $r_1, r_2 \in \mathbb{N}$ such that $r_1 \neq r_2$ and $g^{r_1} = g^{r_2}$. If this was not the case, then g^n would be unique for each $n \in \mathbb{N}$, contradicting the finiteness of G. WLOG take $r_1 < r_2$, and let $r = r_2 - r_1 \in \mathbb{N}$. Using associativity, we then have that:

$$g^{r_1} = g^{r_2} = g^{r_1+r} = g^{r_1}g^r$$

from which we conclude that $g^r = e$. $\square$

Exercise A2.2

Prove Lagrange's Theorem.

Solution

Concepts Involved: Group Axioms, Subgroups, Order, Equivalence Relations.

Let H be a subgroup of a group G and define the relation $\sim$ by $a \sim b$ iff $a = bh$ for some $h \in H$. $\sim$ is reflexive as $a = ae$ (so $a \sim a$) where $e \in H$ is the identity element. $\sim$ is symmetric as if $a \sim b$, then $a = bh$ for some $h \in H$ so $b = ah^{-1}$ (so $b \sim a$) where $h^{-1} \in H$ as H is closed under inverses. Finally $\sim$ is transitive as if $a \sim b$ and $b \sim c$, there exist $h_1, h_2 \in H$ such that $a = bh_1$ and $b = ch_2$ so $a = ch_2h_1$. As $h_2h_1 \in H$ (H is closed under multiplication) it follows that $a \sim c$. Having shown $\sim$ to have these three properties, we conclude it is an equivalence relation. Then, the equivalence classes of $\sim$ partition G, where the equivalence class of $g \in G$ is $[g] = \{gh|h \in H\}$.
Now, let $g \in G$ and define the map $\varphi_g : H \to [g]$ as $\varphi_g(h) = gh$. φ_g is injective as if $\varphi_g(h_1) = \varphi_g(h_2)$ then $gh_1 = gh_2$ and multiplying by g^{-1} on both sides $h_1 = h_2$. φ_g is surjective as if $k \in [g]$, then there exists some $h \in H$ such that $k = gh$ by the definition of $\sim$. Hence φ_g is bijective.
As per our prior observation, the equivalence classes of $\sim$ partition G, so $G = \bigcup_{i=1}^n [g_i]$ and $|G| = \left|\bigcup_{i=1}^n [g_i]\right| = \sum_{i=1}^n \left|[g_i]\right|$. Further, there is a bijection φ_{g_i} from each equivalence class to H, so $\left|[g_i]\right| = |H|$ for all i. Thus $|G| = \sum_{i=1}^n |H| = nH$ and hence $|H|$ divides $|G|$, as desired. $\square$

Exercise A2.3

Show that the order of an element $g \in G$ divides $|G|$.

Solution

Concepts Involved: Group Axioms, Subgroups, Order, Lagrange's Theorem.

Let $g \in G$ with order r. Then, define $H = \{g^n | n \in \mathbb{N}\}$. We claim that H is a subgroup of G. First, $g^n \in G$ for any n as G is closed under multiplicaton, so $H \subset G$. Next, if $g^{n_1}, g^{n_2} \in H$ then $g^{n_1} \cdot g^{n_2} = g^{n_1+n_2} \in H$. Associativity is inherited from the associativity of multiplication in G. Since $g^r = e \in H$, H contains the identity. Finally, for $g^k \in H$ we have $g^{r-k} \in H$ such that $g^k g^{r-k} = g^{r-k} g^k = g^r = e$ so H is closed under inverses. Hence the claim is proven.
Next, we observe that $|H| = r$ as H contains the r elements $e, g, g^2, \ldots g^{r-1}$. Hence by Lagrange's Theorem r divides $|G|$. $\square$

Exercise A2.4

Show that if $y \in G_x$ then $G_y = G_x$

Solution

Concepts Involved: Group Axioms, Conjugacy Classes

Suppose $y \in G_x$. Then there exists some $g \in G$ such that $g^{-1}xg = y$. Multiplying both sides on the left by g and on the right by g^{-1} we find that $x = gyg^{-1}$. We now show the two inclusions.
$\boxed{\subseteq}$ Suppose that $k \in G_x$. Then there exists some $g' \in G$ such that $k = g'^{-1}xg'$. Then using $x = gyg^{-1}$ we find $k = g'^{-1}gyg^{-1}g'$. Now, $g^{-1}g' \in G$ (by closure) and it has inverse $g'^{-1}g$, and hence $k = g'^{-1}gyg^{-1}g' \in G_y$. So, $G_x \subseteq G_y$.
$\boxed{\supseteq}$ Suppose that $l \in G_y$. Then there exists some $g'' \in G$ such that $l = g''^{-1}yg''$. Then with $g^{-1}xg = y$ we find $l = g''^{-1}g^{-1}xgg''$. Much like before, $gg'' \in G$ (by closure) with inverse $g''^{-1}g^{-1}$ so $l \in G_x$. So, $G_y \subseteq G_x$.
We conclude that $G_y = G_x$. $\square$

Exercise A2.5

Show that if x is an element of an Abelian group G, then $G_x = \{x\}$.

Solution

Concepts Involved: Abelian Groups, Conjugacy Classes.

Evidently $x = e^{-1}xe \in G_x$ so $\{x\} \subseteq G_x$. Next, if $k \in G_x$ then $k = g^{-1}xg$ for some $g \in G$, but since G is abelian, $g^{-1}x = xg^{-1}$ so $k = xg^{-1}g = xe = x$ so $k \in \{x\}$ and hence $G_x \subseteq \{x\}$. We conclude that $G_x = \{x\}$. $\square$

Exercise A2.6

Show that any group of prime order is cyclic.

Solution

Concepts Involved: Order, Cyclic Groups.

Suppose $|G| = p$ where p is prime. Since G is finite, every element of G has an order by Exercise A2.1. Since the order of any element $g \in G$ divides $|G| = p$ by Exercise A2.3, and since p is prime, the order of g is either 1 or p. Since $|G| > 1$, there exists at least one $g \in G$ with order p, and this g is a generator of G (with $g^1 = g, g^2, g^3, \ldots, g^p = e$ distinct and comprising all the elements of G. In fact this is true of any non-identity g). Hence G is cyclic. $\square$

Exercise A2.7

Show that every subgroup of a cyclic group is cyclic.

Solution

Concepts Involved: Group Axioms, Subgroups, Cyclic Groups, Euclid's Division Algorithm.

First we prove a necessary Lemma, namely that any nonempty subset of the natural numbers contains a least element. We show this by proving the contrapositive. Suppose that $A \subseteq \mathbb{N}$ has no least element. Then $1 \notin A$ as then 1 would be the least element. Suppose then that $1, \ldots k-1 \notin A$; then $k \notin A$ as then k would be the least element. By strong induction, there exists no $k \in \mathbb{N}$ such that $k \in A$, i.e. A is empty. This concludes the proof of the lemma.
Let $G = \langle a \rangle$ be a cyclic group and H a subgroup of G. If $H = \{e\}$, it is trivially cyclic and we are done. If $H \neq \{e\}$, then there exists some $a^n \in H$ with $n \neq 0$. Since H is closed under inverses, $(a^n)^{-1} = a^{-n} \in H$ as well which ensures that H contains some positive power of a. Then consider the set $A = \left\{k \in \mathbb{N} | a^k \in H\right\}$. Any nonempty subset of the naturals has a minimum element; therefore let $d = \min A$. It is immediate that $\langle a^d \rangle$ is a subgroup of H as $a^d \in H$ and H is a group. To show the reverse containment, suppose that $g \in H$. Since H is a subgroup of the cyclic G, it follows that $g = a^p$ for some $p \in \mathbb{Z}$. We can then write $p = qd + r$ for $0 \leq r < d$ by Euclid's Division algorithm (see Appendix 4). We then have that $a^r = a^{p-qd} = a^p(a^d)^{-q} \in H$ by closure. Now, since d is the least positive integer for which $a^d \in H$ and $0 \leq r < d$, it must follow that $r = 0$. Therefore, $p = qd$ and hence $a^{qd} = (a^d)^q \in \langle a^d \rangle$. So, H is a subgroup of $\langle a^d \rangle$. We conclude that $H = \langle a^d \rangle$ and hence H is cyclic.
$\square$

Exercise A2.8

Show that if $g \in G$ has finite order r, then $g^m = g^n$ if and only if $m = n \pmod r$.

Solution

Concepts Involved: Order, Modular Arithmetic, Euclid's Division Algorithm

Suppose $g \in G$ has finite order r.
$\boxed{\Longleftarrow}$ First suppose that $m = n \pmod r$. Then $m - n = kr$ for some $k \in \mathbb{N}$. Therefore $g^{m-n} = g^{kr}$. But $g^{kr} = (g^r)^k = e^k = e$, so $g^{m-n} = g^m g^{-n} = e$, and multiplying both sides by g^n we find $g^m = g^n$.
$\boxed{\Longrightarrow}$ Suppose $g^m = g^n$. Then multiplying both sides by g^{-n} we find $g^{m-n} = e$. By Euclid's Division algorithm there exist integers q, p such that $m - n = qr + p$ with $0 \leq p < r$. We then have that

$g^{m-n} = g^{qr+p} = g^{qr}g^p = e$. Furthermore, $g^{qr} = (g^r)^q = e^q = e$ so $g^p = e$. But since g has order r and $0 \leq p < r$, it follows that $p = 0$. Hence $m - n = qr$ and so $m \equiv n \pmod r$. $\square$

Exercise A2.9

Cosets define an equivalence relation between elements. Show that $g_1, g_2 \in G$ are in the same coset of H in G if and only if there exists some $h \in H$ such that $g_2 = g_1 h$.

Solution

Concepts Involved: Equivalence Relations, Cosets

In Exercise A2.2 we showed that the relation $\sim$ on a group G defined by $g_1 \sim g_2$ iff $g_1 = g_2 h$ for some $h \in H$ was an equivalence relation. The equivalence classes of this equivalence relation were $\{gh | h \in H\}$, i.e. precisely the left cosets of H in G. So, g_1, g_2 are in the same coset of H in G if and only if $g_1 = g_2 h$ for some $h \in H$, which is exactly what we wished to show. $\square$

Exercise A2.10

How many cosets of H are there in G?

Solution

Concepts Involved: Equivalence Relations, Cosets

We observe that the map $\varphi_g : H \to [g]$ defined in the solution of Exercise A2.2 is a map from H to a right coset of H in G defined by g. Since we showed that this map was bijective, this shows that $|H| = |Hg|$ for any $g \in G$. Furthermore, since the cosets define an equivalence relation between elements of G, the cosets of H in G partition G. So, we conclude that there are $|G|/|H|$ cosets of H in G, each of cardinality $|H|$. $\square$

Exercise A2.11: Characters

Prove the properties of characters given above.

Solution

Concepts Involved: Matrix Groups, Character (Trace)

Recall that the character of a matrix group $G \subset M_n$ is a function on the group defined by $\chi(g) = \mathrm{tr}(g)$ where tr is the trace function. It has the properties that (1) $\chi(I) = n$, (2) $|\chi(g)| \leq n$, (3) $|\chi(g)| = n$ implies $g = e^{i\theta}I$, (4) χ is constant on any given conjugacy class of G, (5) $\chi(g^{-1}) = \chi^*(g)$ and (6) $\chi(g)$ is an algebraic number for all g.

The six properties are proven below.

(1) $\chi(I) = \mathrm{tr}(I) = \sum_{k=1}^{n} 1 = n$.

(2) Let $g \in G$. Since G is finite, by Exercise A2.1 it follows that g has order r such that $g^r = I$. So, g may be diagonalized with roots of unity $e^{2\pi i j/r}$, $j \in \{0, 1, \ldots, r-1\}$ on the diagonal. We then

find using the triangle inequality that:

$$\left|\chi(g)\right| = \left|\operatorname{tr}(g)\right| = \left|\sum_{k=1}^{n} e^{2\pi i j_k/r}\right| \leq \sum_{k=1}^{n} \left|e^{2\pi i j_k/r}\right| = \sum_i 1 = n$$

which proves the claim.

(3) The (complex) triangle inequality $|z_1 + z_2| \leq |z_1| + |z_2|$ is saturated when $z_1 = kz_2$ for some $k \geq 0$. This can only occur in the above equation when every λ_i in the sum is identical (as distinct roots of unity are not related by a non-negative constant). If the λ_is are identical, then g is diagonal with diagonal entries of unit modulus, so $g = e^{i\theta}I$ as claimed.

(4) Let $G_x = \{g^{-1}xg | g \in G\}$ be the conjugacy class of x in G. We then have for any $h \in G_x$ that $\chi(h) = \chi(g^{-1}xg) = \operatorname{tr}(g^{-1}xg) = \operatorname{tr}(xgg^{-1}) = \operatorname{tr}(xI) = \operatorname{tr}(x)$, using the cyclicity of the trace. We conclude that χ is constant on the conjugacy class.

(5) By the same argument as (2), $g \in G$ can be diagonalized with roots of unity $e^{2\pi i j/r}$ on the diagonal:

$$g = \begin{bmatrix} e^{2\pi i j_1/r} & & & \\ & e^{2\pi i j_2/r} & & \\ & & \ddots & \\ & & & e^{2\pi i j_n/r} \end{bmatrix}$$

It then follows that g^{-1} is:

$$g^{-1} = \begin{bmatrix} e^{-2\pi i j_1/r} & & & \\ & e^{-2\pi i j_2/r} & & \\ & & \ddots & \\ & & & e^{-2\pi i j_n/r} \end{bmatrix}$$

So we have that:

$$\chi(g^{-1}) = \operatorname{tr}\left(g^{-1}\right) = \sum_{j=1}^{n} e^{-2\pi i j_k/r} = \sum_{j=1}^{n} (e^{2\pi i j_k/r})^* = \left(\sum_{j=1}^{n} e^{2\pi i j_k/r}\right)^* = (\operatorname{tr}(g))^* = \chi^*(g).$$

which proves the claim.

(6) $\chi(g)$ is the sum of r-th roots of unity, which are algebraic; hence $\chi(g)$ is algebraic as the sum of algebraic numbers.

□

Exercise A2.12: Unitary matrix groups

$(*)$ A unitary matrix group is comprised solely of unitary matrices (those who which satisfy $U^\dagger U = I$). Show that every matrix group is equivalent to a unitary matrix group. If a representation of a group consists entirely of unitary matrices, we may refer to it as being a *unitary representation*.

Solution

Concepts Involved: Matrix Groups, Character (Trace), Equivalence, Unitary Operators.
Recall that two groups are equivalent if they are isomorphic (i.e. there is a bijection between the groups that respects the group multiplication) and the isomorphic element have the same character.

Let $G = \{A_1, \dots A_n\}$ be a finite matrix group. Then define

$$A = \sum_{i=1}^{n} A_i^\dagger A_i.$$

By Ex. 2.25 each term of the above sum is positive, and by Ex. 2.24 each term is Hermitian. The sum of Hermitian operators is Hermitian, so A is Hermitian. By Ex. 2.21, A is diagonalizable. Let U be the unitary matrix that diagonalizes A. We then have that D is a diagonal matrix, with:

$$D = UAU^\dagger.$$

Let $D^{1/2}$ be the matrix obtained by taking the square root of the diagonal entries of D. Then define $T = D^{1/2}U$. We then claim that $G_U = \{V_1, \dots V_n\}$ is a unitary matrix group equivalent to G, where:

$$V_i = TA_iT^{-1}.$$

We have three points to verify; (i) That the V_is are unitary, (ii) That $\varphi : G \to G_u$ defined by $\varphi(A_i) = TA_iT^{-1} = V_i$ is an isomorphism, and (iii) that the characters of A_i and V_i are equivalent.

(i) For any V_i, we have:

$$\begin{aligned}
V_i^\dagger V_i &= (TA_iT^{-1})^\dagger(TA_iT^{-1}) \\
&= (D^{1/2}UA_iU^\dagger D^{-1/2})^\dagger(D^{1/2}UA_iU^\dagger D^{-1/2}) \\
&= (D^{-1/2}UA_i^\dagger U^\dagger D^{1/2})(D^{1/2}UA_iU^\dagger D^{-1/2}) \\
&= (D^{-1/2}UA_i^\dagger U^\dagger)D(UA_iU^\dagger D^{-1/2}) \\
&= (D^{-1/2}UA_i^\dagger U^\dagger)(UAU^\dagger)(UA_iU^\dagger D^{-1/2}) \\
&= (D^{-1/2}UA_i^\dagger)A(A_i^\dagger U^\dagger D^{-1/2}) \\
&= (D^{-1/2}UA_i^\dagger)\left(\sum_{j=1}^{n} A_j^\dagger A_j\right)(A_iU^\dagger D^{-1/2}) \\
&= D^{-1/2}U\left(\sum_{j=1}^{n}(A_jA_i)^\dagger(A_jA_i)\right)U^\dagger D^{-1/2} \\
&= D^{-1/2}U\left(\sum_{k=1}^{n} A_k^\dagger A_k\right)U^\dagger D^{-1/2} \\
&= D^{-1/2}UAU^\dagger D^{-1/2} \\
&= D^{-1/2}DD^{-1/2} \\
&= I
\end{aligned}$$

Where in the sixth equality we use the unitarity of U, and in the ninth equality we use that $A_j A_i = A_k$ iterates over all the group elements as A_j iterates over all the group elements. To see that this is the case, it suffices to show that the map $\psi_i : M_n \to M_n$ defined by $\psi_i(A_j) = A_j A_i$ is a bijection. To see that it is injective, suppose that $\psi_i(A_{j_1}) = \psi_i(A_{j_2})$. Then it follows that $A_{j_1} A_i = A_{j_2} A_i$, and multiplying on the left by A_i^{-1} (which exists) we find that $A_{j_1} = A_{j_2}$. To see that it is surjective, suppose that $A_{j'} \in M_n$. Then, there exists $A_{j'} A_i^{-1} \in M_n$ such that $\psi_i(A_{j'} A_i^{-1}) = A_{j'} A_i^{-1} A_i = A_{j'}$. We conclude that ψ_i is bijective.

(ii) Firstly, φ is a homomorphism as for any A_i, A_j we have:

$$\varphi(A_i)\varphi(A_j) = V_i V_j = TA_iT^{-1}TA_jT^{-1} = TA_iA_jT^{-1} = \varphi(A_iA_j).$$

Next, φ is surjective by construction. Finally, it is injective; suppose that $V_i = V_j$. Then we have that:

$$TA_iT^{-1} = TA_jT^{-1}$$

And multiplying both sides on the left by T^{-1} on the left and T on the right we find that $A_i = A_j$. Hence we conclude that φ is a bijective homomorphism and hence an isomorphism.

(iii) This is immediate from the cyclicity of the trace:

$$\chi(V_i) = \mathrm{tr}\Big(TA_iT^{-1}\Big) = \mathrm{tr}\Big(T^{-1}TA_i\Big) = \mathrm{tr}(A_i) = \chi(A_i).$$

The claim is therefore proven. □

Exercise A2.13

Show that every irreducible Abelian matrix group is one dimensional.

Exercise A2.14

Show that if ρ is an irreducible representation of G, then $|G|/d_\rho$ is an integer.

Exercise A2.15

Using the Fundamental Theorem, prove that characters are orthogonal, that is:

$$\sum_{i=1}^{r} r_i(\chi_i^p)^*\chi_i^q = |G|\delta_{pq} \text{ and } \sum_{p=1}^{r}(\chi_i^p)^*\chi_j^q = \frac{|G|}{r_i}\delta_{ij}$$

where p, q, and δ_{pq} have the same meaning as in the theorem and χ_i^p is the value the character of the pth irreducible representation takes on the ith conjugacy class of G and r_i is the size of the ith conjugacy class of G and r_i is the size of the ith conjugacy class.

Exercise A2.16

S_3 is the group of permutations of three elements. Suppose we order these as mapping 123 to: 123;231;312;213;132, and 321, respectively. Show that there exist two one-dimensional irreducible representations of S_3, one of which is trivial, and the other of which is 1,1,1,-1,-1,-1, corresponding in order to the six permutations given earlier. Also show that there exists a two dimensional irreducible representation, with the matrices

$$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \quad \frac{1}{2}\begin{bmatrix}-1 & -\sqrt{3}\\ \sqrt{3} & -1\end{bmatrix}, \quad \frac{1}{2}\begin{bmatrix}-1 & \sqrt{3}\\ -\sqrt{3} & 1\end{bmatrix},$$
$$\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}, \quad \frac{1}{2}\begin{bmatrix}1 & \sqrt{3}\\ \sqrt{3} & -1\end{bmatrix}, \quad \frac{1}{2}\begin{bmatrix}1 & -\sqrt{3}\\ -\sqrt{3} & 1\end{bmatrix}$$

Verify that the representations are orthogonal.

Exercise A2.17

Prove that the regular representation is faithful.

Exercise A2.18

Show that the character of the regular representation is zero except on the representation of the identity element, for which $\chi(I) = |G|$.

Exercise A2.19

Use Theorem A2.5 to show that the regular representation contains d_{ρ^p} instances of each irreducible representation ρ^p. Thus, if R denotes the regular representation, and $\hat{G}$ denotes the set of all inequivalent irreducible representations, then:

$$\chi_i^R = \sum_{\rho \in G} d_\rho \chi_i^\rho$$

Exercise A2.20

The character of the regular representation is zero except for the conjugacy class i containing e, the identity element in G. Show, therefore, that

$$\sum_{\rho \in G} d_\rho \chi^\rho(g) = N\delta_{ge}.$$

Exercise A2.21

Show that $\sum_{\rho \in \hat{G}} d_\rho^2 = |G|$.

Exercise A2.22

Substitute (A2.10) into (A2.9) and prove that $\hat{f}(\rho)$ is obtained.

Exercise A2.23

Let us represent an Abelian group G by $g \in [0, N-1]$, with addition as the group operation, and define $\rho_h(g) = \exp\left[-2\pi igh/N\right]$ at the h representation of g. This representation is one-dimensional, so $d_\rho = 1$. Show that the Fourier transform relations for G are

$$\hat{f}(h) = \frac{1}{\sqrt{N}} \sum_{g=0}^{N-1} f(g) e^{-2\pi igh/N} \text{ and } f(h) = \frac{1}{\sqrt{N}} \sum_{g=0}^{N-1} \hat{f}(g) e^{2\pi igh/N}$$

Exercise A2.24

Using the results of Exercise A2.16, construct the Fourier transform over S_3 and express it as a 6x6 unitary matrix.